Notes on Lévy Processes

P. Ouwehand

Department of Mathematical Sciences
Stellenbosch University

Contents

1 Definition of Lévy Processes

Throughout this section, we work with a stochastic base $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$, where the filtration $(\mathcal{F}_t)_t$ satisfies the usual conditions.

Definition 1.1 A $\mathbb{R}^d$–valued stochastic process $X = (X_t)_{t\geq 0}$ is called a *Lévy process* iff the following conditions hold:

(i) *Stationary Increments:* $X_{t+h} - X_t \stackrel{\mathrm{d}}{=} X_h$ (for $h \geq 0$).

(ii) *Independent Increments:* $X_{t+h} - X_t$ is independent of $\mathcal{F}_t$ (for $h \geq 0$).

(iii) *Stochastic Continuity:* $X_s \to X_t$ in probability as $s \to t$.

(iv) *Càdlàg:* Almost surely, sample paths of X have left limits and are right–continuous.

□

Observe that, by (i), we have $X_0 = 0$ a.s. For the moment, note that (arithmetic) Brownian motion is a Lévy process. A Poisson process is another example, with very different properties. We will give more examples of Lévy processes later on.

Remarks 1.2 A process which satisfies only (i)-(iii) of Defn 1.1 is called a Lévy process *in law*. The difference between Lévy processes and Lévy processes in law is slight: It can be shown that any Lévy process in law has a modification which is càdlàg: see Thm B.2 for a proof.

□

2 Distributions Associated with Lévy Processes

The fact that a Lévy process has independent stationary increments implies some very strong constraints on the set possible distributions of X_t. Note that, for any $n \in \mathbb{N}$, we can write

$$X_t = \sum_{i=1}^{n} Y_i \qquad \text{where} \qquad Y_i = X_{ti/n} - X_{t(i-1)/n}$$

Clearly the Y_i are i.i.d. variables. Random variables with this property are called infinitely divisible, and we study their properties next.

2.1 Convolution

Suppose that X, Y are independent $\mathbb{R}^d$–valued RVs with distributions μ, ν respectively. The distribution of the sum $X + Y$ is given by the *convolution* $\mu * \nu$ of the measures: For a Borel set $B \in \mathcal{B}(\mathbb{R}^d)$ we have

$$\begin{aligned}\mathbb{P}(X + Y \in B) &= \int_{\mathbb{R}^d} \mathbb{P}(X \in B - y|Y = y)\ \nu(dy) \\ &= \int_{\mathbb{R}^d} \mu(B - y)\ \nu(dy) \\ &= \int_{\mathbb{R}^d \times \mathbb{R}^d} I_B(x + y)\ (\mu \times \nu)(d(x, y))\end{aligned}$$

Definition 2.1 Let μ, ν be probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. We define the *convolution* of μ, ν to be a set function $\mu * \nu : \mathcal{B}(\mathbb{R}^d) \to \mathbb{R}$ given by

$$\mu * \nu(B) = \int_{\mathbb{R}^d \times \mathbb{R}^d} I_B(x + y)\ (\mu \times \nu)(d(x, y))$$

□

It is easy to verify that $\mu * \nu$ is again a probability measure on $\mathbb{R}^d$. We leave the following as an exercise:

Proposition 2.2 *The set $\mathcal{M}_1(\mathbb{R}^d)$ of probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)$ forms a commutative monoid when equipped with the convolution operation. That is:*

(i) $*$ is a commutative and associative operation on $\mathcal{M}_1(\mathbb{R}^d)$:

$$\mu * \nu = \nu * \mu \qquad \mu * (\nu * \gamma) = (\mu * \nu) * \gamma \qquad \text{for all } \mu, \nu, \gamma \in \mathcal{M}_1(\mathbb{R}^d)$$

(ii) The Dirac measure δ_0 (unit mass at the origin) is an identity element for $*$:

$$\mu * \delta_0 = \mu = \delta_0 * \mu \qquad \text{for all } \mu \in \mathcal{M}_1(\mathbb{R}^d)$$

□

A standard argument (starting from indicator functions) shows that

$$\int_{\mathbb{R}^d} f(x)\ \mu * \nu(dx) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x + y)\ \mu(dx)\ \nu(dy)$$

whenever f is a bounded measurable function.

2.2 Characteristic Functions

Recall that the characteristic function (or Fourier transform) of a probability measure μ on $\mathbb{R}^d$ is a map $\hat{\mu} : \mathbb{R}^d \to \mathbb{C}$ defined as follows:

$$\hat{\mu}(u) = \int_{\mathbb{R}^d} e^{i\langle u,x\rangle}\ \mu(dx)$$

The characteristic function φ_X of a random variable X is defined to be the characteristic function of its distribution $\mathbb{P}_X$, i.e.

$$\varphi_X(u) = \hat{\mathbb{P}}_X(u) = \mathbb{E}[e^{i\langle u,X\rangle}]$$

It is easy to verify that

$$\widehat{\mu * \nu}(u) = \hat{\mu}(u)\hat{\nu}(u) \qquad \hat{\delta}_0(u) = 1$$

(i.e. that the Fourier transform is a homomorphism from the monoid $\mathcal{M}_1(\mathbb{R}^d)$ with convolution to the monoid of bounded continuous $\mathbb{C}$–valued functions on $\mathbb{R}^d$ with ordinary multiplication.)

We collect here (following Sato[**?**]) several well–known results about characteristic functions. (See, e.g., Billingsley[**?**] or Kallenberg[**?**] for proofs.)

Theorem 2.3 μ, ν, μ_n *are probability distributions on* $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$*.* X, Y, X_n *are* $\mathbb{R}^d$*–valued random variables.*

(a) (Uniqueness Theorem): If $\hat{\mu} = \hat{\nu}$*, then* $\mu = \nu$*.*

(b) (Bochner's Theorem ⇒): $\hat{\mu}$ *is a uniformly continuous function, with* $\hat{\mu}(0) = 1, |\hat{\mu}(u)| \leq 1$*. Furthermore,* $\hat{\mu}$ *is* non–negative definite*, i.e. for any* $u_1, \ldots, u_n \in \mathbb{R}^d$ *and* $\xi_1, \ldots, \xi_n \in \mathbb{C}$ *we have*

$$\sum_{j=1}^{n}\sum_{k=1}^{n} \hat{\mu}(u_j - u_k)\xi_j\bar{\xi}_k \geq 0$$

(c) (Bochner's Theorem ⇐): Conversely if φ *is a map* $\varphi : \mathbb{R}^d \to \mathbb{C}$ *which has* $\varphi(0) = 1$*, is continuous at* $u = 0$ *and is non–negative definite, then* φ *is the characteristic function of some probability distribution on* $\mathbb{R}^d$*.*

(d) (Glivenko's Theorem): If $\mu_n \to \mu$ *weakly (i.e., in distribution) then* $\hat{\mu}_n \to \hat{\mu}$ *uniformly on compacts. Conversely, if* $\hat{\mu}_n \to \hat{\mu}$ *pointwise, then* $\mu_n \to \mu$ *weakly.*

(e) (Lévy Continuity Theorem): If $\mu_n \to \varphi$ *pointwise, and if* φ *is continuous at* $u = 0$*, then* φ *is the characteristic function of a distribution on* $\mathbb{R}^d$*.*

(f) (Kac's Theorem): Let $X = (X_1, \ldots X_n)$ *be an* $\mathbb{R}^{nd}$*–valued RV. Then* $X_1, \ldots, X_n$ *are independent iff*

$$\hat{\mathbb{P}}_X(z) = \prod_{j=1}^{n} \hat{\mathbb{P}}_{X_j}(z_j) \qquad \textit{for } z = (z_1, \ldots, z_n), \textit{ where } z_j \in \mathbb{R}^d$$

(g) If X, Y *are independent* $\mathbb{R}^d$*–valued RVs, then* $\hat{\mathbb{P}}_{X+Y}(u) = \hat{\mathbb{P}}_X(u)\hat{\mathbb{P}}_Y(u)$*.*

(h) (Moments): Let $n \in \mathbb{N}$*. If* μ *has absolute moment of order* n *(i.e. if* $\int |x|^n \, \mu(dx) < \infty$*), then* $\hat{\mu}$ *is a* C^n*–function, and*

$$\int x_1^{n_1} \cdot \cdots \cdot x_d^{n_d} \mu(dx) = \left[\left(\frac{1}{i} \frac{\partial}{\partial u_1} \right)^{n_1} \cdots \left(\frac{1}{i} \frac{\partial}{\partial u_d} \right)^{n_d} \right] \Big|_{u=0} \hat{\mu}(u)$$

for any non–negative integers $n_1, \ldots n_d$ *satisfying* $n_1 + \cdots + n_d \leq n$*.*

□

2.3 Infinite Divisibility

Given $\mu \in \mathcal{M}_1(\mathbb{R}^d)$ and $n \in \mathbb{N}$, define the n–fold convolution of μ by

$$\mu^n = \mu * \mu * \cdots * \mu \qquad (n \text{ times}), \qquad \mu^0 = \delta_0$$

Definition 2.4 A probability μ distribution on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)$ is called infinitely divisible iff it has an n^{th} convolution root $\nu = \mu^{\frac{1}{n}}$ for every $n \in \mathbb{N}$, i.e. iff for every $n \in \mathbb{N}$ there is a probability measure ν such that $\mu = \nu^n$.
A $\mathbb{R}^d$–valued RV is said to be infinitely divisible iff its distribution $\mathbb{P}_X$ is.

□

Note that μ is infinitely divisible iff, for each $n \in \mathbb{N}$, $\hat{\mu}$ has a n^{th} root which is the characteristic function of a distribution: Indeed, if $\mu = \nu^n$, then $\hat{\mu} = \hat{\nu}^n$. This simple observation allows us, in some cases, to determine whether a distribution is infinitely divisible merely by inspecting its characteristic function.

Examples 2.5 (a) Let μ be a normal distribution with mean vector $\gamma \in \mathbb{R}^d$ and covariance matrix $\Sigma \in \mathbb{R}^{d \times d}$. The characteristic function is given by

$$\hat{\mu}(u) = \exp(i\langle u, \gamma\rangle - \tfrac{1}{2}\langle u, \Sigma u\rangle)$$

Clearly,

$$\exp\left(\tfrac{1}{n}(i\langle u, \gamma\rangle - \langle u, \Sigma u\rangle)\right)$$

is an n^{th} root of $\hat{\mu}$ which is also the characteristic function of a normal distribution with mean $\frac{1}{n}\gamma$ and covariance matrix $\frac{1}{n}\Sigma$. Hence a normal distributed RV is infinitely divisible.

(b) On $\mathbb{R}^1$, let μ be a Poisson distribution with mean λ (i.e. μ is supported by the non–negative integers and has $\mu\{k\} = e^{-\lambda}\frac{\lambda^k}{k!}$ for $k = 0, 1, 2, \ldots$) The characteristic function is given by

$$\hat{\mu}(u) = \exp(\lambda(e^{iu} - 1))$$

Clearly,

$$\exp(\tfrac{\lambda}{n}(e^{iu} - 1))$$

is an n^{th} root of $\hat{\mu}$ which is also the characteristic function of a Poisson distribution with mean $\frac{\lambda}{n}$. Hence a Poisson–distributed RV is infinitely divisible.

□

Other examples of infinitely divisible distributions are the exponential–, Γ–, geometric, and Cauchy distributions, and this follows immediately by inspecting their characteristic functions (which may be found in Sato[**?**]). There are many other distributions where proving infinite divisibility is much more difficult, e.g. the Student's–t, Pareto, lognormal, Gumbel, Weibull and F-distributions (see Sato[**?**], section 8, for references).

Proposition 2.6 *(a) The convolution of two infinitely divisible distributions is infinitely divisible.*

(b) The sum of two independent infinitely divisible RVs is infinitely divisible.

Proof: (a) If $\mu_i = \nu_i^n$ $(i = 1, 2)$, then $\mu_1 * \mu_2 = (\nu_1 * \nu_2)^n$.
(b) follows immediately from (a) and the fact that if X, Y are independent, then $\mathbb{P}_{X+Y} = \mathbb{P}_X * \mathbb{P}_Y$.

$\dashv$

The proof of the next theorem may be found in Appendix A

Theorem 2.7 *(a) If μ is an infinitely divisible distribution on $\mathbb{R}^d$, then there is a unique continuous function $f : \mathbb{R}^d \to \mathbb{C}$ such that $f(0) = 0$ and $\hat{\mu}(u) = e^{f(u)}$.*
We call f the distinguished logarithm *of $\hat{\mu}$, and write $f(u) = \log \hat{\mu}(u)$.*

(b) If μ is infinitely divisible, then its n^{th} convolution root is unique, i.e. there is a unique *distribution ν such that $\mu = \nu^n$.*

(c) If μ_n $(n \in \mathbb{N})$ are infinitely divisible, and if $\mu_n \to \mu$, then μ is infinitely divisible.

□

Recall that if μ is infinitely divisible, then $\mu^{1/n}$ is simply its n^{th} convolution root, and this convolution root is unique, by the preceding theorem. We may therefore define $\mu^{p/q}$ for rational $p/q > 0$ by $\mu^{p/q} = (\mu^{1/q})^p =$ the p–fold convolution of $\mu^{1/q}$. We also define $\mu^0 = \delta_0$. We can extend the definition of convolution powers to arbitrary non–negative reals as follows:

Proposition 2.8 *If μ is infinitely divisible, then we may define μ^t for any $t \geq 0$. Then μ^t is infinitely divisible.*

Proof: Suppose that μ is infinitely divisible. We already have defined $\mu^{\frac{1}{n}}$ for $n \in \mathbb{N}$. Then $\mu^{\frac{1}{n}}$ is clearly itself infinitely divisible, as $\hat{\mu}(u)^{\frac{1}{n}} = (\hat{\mu}(u)\frac{1}{nk})^k$ for any $k \in \mathbb{N}$. $\hat{\mu}(u)^{\frac{m}{n}}(u) = (\hat{\mu}(u)\frac{1}{n})^m$ shows that we may define an infinitely divisible distribution μ^q for every non–negative rational number q. If t is irrational, choose a sequence of rationals $q_n \to t$. Then $\hat{\mu}(u)^{q_n} \to \hat{\mu}(u)^t$. Since $\hat{\mu}(u)^t$ is continuous, it is a characteristic function of some probability distribution (by Thm. 2.3 (Lévy continuity)). We define μ^t to be that distribution. Then $\mu^{q_n} \to \mu^t$ weakly, and hence μ^t is infinitely divisible, by Thm. 2.7.

$\dashv$

2.4 Lévy Processes and Infinite Divisibility

For Lévy processes, the following result is important:

Proposition 2.9 *Let X be a Lévy process in $\mathbb{R}^d$. If $t \geq 0$, the law of X_t is an infinitely divisible distribution on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Moreover, if μ is the law of X_1, then μ^t is the law of X_t.*

Proof: Let $t \geq 0$ and $n \in \mathbb{N}$. Then

$$X_t = (X_{t/n} - X_0) + (X_{2t/n} - X_{t/n}) + \cdots + (X_t - X_{t(n-1)/n})$$

represents X_t as the sum of independent identically distributed random variables. Thus if μ_t is the law of X_t, then clearly $\mu_t = (\mu_{t/n})^n$, which establishes the infinite divisibility of μ_t. To see that $\mu_t = \mu_1^t$, first note that this holds if t is rational: If $t = \frac{p}{q} \geq 0$, then $X_{p/q}$ is the sum of p–many independent random variables with distribution $\mu_{1/q}$. For general $t \geq 0$, the result follows by Propn. 2.8.

$\dashv$

We quote here an important technical result, whose proof may be found in Appendix B.

Theorem 2.10 *Let μ be an infinitely divisible distribution on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Then there is a Lévy process X such that μ is the distribution of X_1.*
Moreover, if X' is another such Lévy process, then X, X' are identical in law.

$\square$

Thus Lévy processes and infinitely divisible distributions are intimately related: If X is a Lévy process, then each X_t is infinitely divisible. Conversely, if μ is an infinitely divisible distribution there is a Lévy process X such that $X_t = \mu^t$, i.e. the law of X is determined by μ. The study of Lévy processes is therefore very much the study of infinitely divisible distributions.

We now take a first look at the characteristic functions associated with a Lévy process. These are fully characterized by the Lévy–Ito Decomposition Theorem, to be proved later. First note that if μ is an infinitely divisible distribution, then there is a unique continuous function $\eta : \mathbb{R}^d \to \mathbb{C}$ satisfying $\eta(0) = 0$ and $\hat{\mu}(u) = e^{\eta(u)}$: This is just Thm. 2.7(a). Suppose now that μ is the law of X_1, where X is a Lévy process. It follows immediately that X_t has law μ^t, and thus that the characteristic function of X_t is given by

$$\mathbb{E}[e^{i\langle u, X_t\rangle}] = e^{t\eta(u)}$$

The function η is known as the Lévy exponent (or characteristic exponent, or Lévy symbol) of the process X. Clearly it completely determines the law of X.

3 The Strong Markov Property

We will prove in this section that every Lévy process is a strong Markov process. The proof follows Protter[**?**]. First, we will need the following easy result, which will also be useful in the proof of the Lévy–Ito decomposition theorem:

Proposition 3.1 *Suppose that X is a Lévy process in $\mathbb{R}^d$, and that $u \in \mathbb{R}^d$. Then*

$$M_t^u := \frac{e^{i\langle u, X_t\rangle}}{\mathbb{E}[e^{i\langle u, X_t\rangle}]}$$

is a càdlàg square–integrable martingale.

Proof: Let η be the Lévy exponent of X. Clearly

$$\begin{aligned}
\mathbb{E}[M_t^u|\mathcal{F}_s] &= e^{-t\eta(u)}\mathbb{E}[e^{i\langle u, X_t\rangle}|\mathcal{F}_s] \\
&= e^{-t\eta(u)}e^{i\langle u, X_s\rangle}\mathbb{E}[e^{i\langle u, X_t - X_s\rangle}|\mathcal{F}_s] \\
&= e^{-t\eta(u)}e^{i\langle u, X_s\rangle}\mathbb{E}[e^{i\langle u, X_t - X_s\rangle}] \\
&= e^{-t\eta(u)}e^{i\langle u, X_s\rangle}e^{(t-s)\eta(u)} \\
&= M_s^u
\end{aligned}$$

This establishes the martingale property. That M_t^u is càdlàg follows because X is càdlàg . The square–integrability of M_t^u is follows by a simple calculation:

$$\mathbb{E}[|M_t^u|^2] = |e^{-2t\eta(u)}|\, \mathbb{E}[|e^{i\langle u, X_t\rangle}|^2] = |e^{-t\eta(u)}|^2 < \infty$$

$\dashv$

Theorem 3.2 *Let X be a Lévy process on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0})$, and let τ be a stopping time. On the set $\{\tau < \infty\}$, the process*

$$Y_t := X_{\tau+t} - X_\tau$$

is a Lévy process adapted to $\mathcal{G}_t = \mathcal{F}_{\tau+t}$. Furthermore, Y is independent of $\mathcal{F}_\tau$ and has the same law as X.

Proof: Suppose first of all that τ is bounded. Let $\{u_j : j \in \mathbb{N}\}$ enumerate a countable dense subset of $\mathbb{R}^d$ (e.g. $\mathbb{Q}^d$), and let $(t_j : j \in \mathbb{N})$ be an increasing sequence in $\mathbb{R}_+$. Let $\varphi_t(u) := \mathbb{E}[e^{i\langle u, X_t\rangle}]$ be the characteristic function of X_t. Then $M_t^u = \frac{e^{i\langle u, X_t\rangle}}{\varphi_t(u)}$ is a martingale, for each u. Thus if $A \in \mathcal{F}_\tau$, we have

$$\begin{aligned}\mathbb{E}\left[I_A e^{i\sum_{j=1}^n \langle u_j, X_{\tau+t_j} - X_{\tau+t_{j-1}}\rangle}\right] &= \mathbb{E}\left[I_A \prod_{j=1}^n \frac{M^{u_j}_{\tau+t_j}}{M^{u_j}_{\tau+t_{j-1}}} \frac{\varphi_{\tau+t_j}(u_j)}{\varphi_{\tau+t_{j-1}}(u_j)}\right] \\ &= \mathbb{P}(A) \prod_{j=1}^n \varphi_{t_j - t_{j-1}}(u_j)\end{aligned}$$

using the optional sampling theorem to condition successively on $\tau + t_{n-1}, \tau + t_{n-2}, \ldots, \tau$, and the fact that $\varphi_{\tau(\omega)+t_j} = \varphi_{\tau(\omega)}\varphi_{t_j}$. It follows immediately that $\mathcal{F}_\tau$ is independent of Y, and also that Y has stationary independent increments with the same law as X.

If τ is not bounded, the above remains true when applied to the bounded stopping times $\tau \wedge n$ and the set $A \cap \{\tau \leq n\}, A \in \mathcal{F}_\tau$. Allowing $n \to \infty$ and applying the dominated convergence theorem shows that the result hods for unbounded τ and events $A \cap \{\tau < \infty\}, A \in \mathcal{F}_\tau$.

$\dashv$

4 Poisson Process; Compound Poisson Process

Here is a useful lemma, taken from Feller[**?**]:

Lemma 4.1 *Suppose that $u : \mathbb{R}_+ \to \mathbb{R}_+$ is a strictly positive function satisfying*

$$u(t+s) = u(t)u(s)$$

and that u is bounded on compact sets. Then $u(t) = e^{\lambda t}$ for some $\lambda \in \mathbb{R}$

Proof: Let $\lambda = \ln u(1)$, and define $v(t) = e^{-\lambda t}u(t)$. It suffices to show that $v \equiv 1$.

Certainly $v(1) = 1$. Hence if $q \in \mathbb{N}$, then $1 = v(1) = v(\frac{q}{q}) = v(\frac{1}{q})^q$, from which it follows that $v(\frac{1}{q}) = 1$. It follows that if $\frac{p}{q} \geq 0$ is rational (where $q > 0$), then $v(\frac{p}{q}) = v(\frac{1}{q})^p = 1$ also.

Now suppose that $v(t) \neq 1$ for some $t \geq 0$. Such a t cannot be rational. If $v(t) = c > 0$, then $v(nt) = c^n \to \infty$. So v takes on arbitrarily large values. Similarly, if $v(t) = c < 1$, choose $N \in \mathbb{N}$ such that $N > na$, and note that $v(N - na) = c^{-n} \to \infty$. Hence v takes on arbitrarily large values in this case also. Let L be such a large value. If $v(t) = L$, then $v(t - \frac{p}{q}) = L$ also, so that v takes on the value L on all intervals.

Thus, if $v \not\equiv 1$, then v is unbounded on compact intervals, and hence so is u, contradiction.

⊣

Let $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F} := (\mathcal{F}_t)_{t\geq 0})$ be a stochastic base. Let $(T_n : n \in \mathbb{N})$ be a strictly increasing sequence of random times (i.e. non–negative random variables), and let $T_0 = 0$ a.s.. Associated with such a sequence is a *counting process*, i.e. a process N_t with values in $\mathbb{N} \cup \{\infty\}$ defined by

$$N_t = \sum_{n\geq 1} I_{\{t \geq T_n\}}$$

Loosely, if T_n is the time of the n^{th} event, then N_t is the number of events that have occurred by time t. It is easy to show that N_t is adapted to $\mathbb{F}$ iff the T_n are $\mathbb{F}$–stopping times, because $\{N_t = n\} = \{T_n \leq t < T_{n+1}\}$, and $\{T_n \leq t\} = \{N_t \geq n\}$.

Definition 4.2 *A* Poisson process *is an adapted counting process which is also a Lévy process.*

Theorem 4.3 *Let N be a Poisson process. Then there exists $\lambda > 0$ such that*

$$\mathbb{P}(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$$

i.e. N_t has a Poisson distribution with mean λt.
Moreover, if N_t is a counting process for the sequence $(\tau_n)_{n\geq 1}$ of stopping times, then the collection of interarrival times $\{\tau_n - \tau_{n-1} : n \geq 1\}$ is a set of independent exponentially distributed random variables with parameter λ.

Proof: Let $u(t) = \mathbb{P}(N_t = 0)$. Since N_t is continuous in probability (being a Lévy process), the function $u(t)$ is a decreasing non–negative continuous function. Clearly, if $0 \leq s < t$, then $\{N_t = 0\} = \{N_s = 0\} \cap \{N_t - N_s = 0\}$, and hence $u(t) = u(s)u(t - s)$, by the fact that N_t has stationary independent increments.

We now claim that u is strictly positive. For suppose that $u(t) = 0$ for some t. Certainly $u(s) = 0$ for all $s \geq t$. Given $0 < s < t$, choose $n \in \mathbb{N}$ such that $ns > t$. Then $0 = u(ns) = u(s)^n$. Hence $u(s)$ is zero for all $0 < s < t$ also. But then $\tau_1 \leq t$ a.s. for all t, which is impossible. Hence u is strictly positive.

By lemma 4.1, $u(t) = e^{-\lambda t}$ for some $\lambda \geq 0$. If $u(t) = 0$ for some t, then the continuity of u would imply that $u \equiv 0$ identically. For, given $s < t$, choose $n \in \mathbb{N}$ such that $ns > t$. Then $0 = u(ns) = u(s)^n$. Hence $u(s)$ is zero for all $s < t$ also. But this is impossible, since N is a counting process.) Since also $\{\tau_1 > t\} = \{N_t = 0\}$ we see that τ_1 is exponentially distributed, with parameter λ.

Now define $N_t^n = N_{\tau_n + t} - N_{\tau_n}$. Then clearly $N_t^n = \sum_{m=n+1}^{\infty} I_{\{t \geq \tau_m - \tau_n\}}$ is a counting process for the strictly increasing sequence of random times $\{\tau_m - \tau_n : m > n\}$. By the strong Markov property, N_t^n is a Lévy process (adapted to $\mathcal{F}_{\tau_n + t}$) with the same law as N_t, and thus $\tau_{n+1} - \tau_n$ is independent of $\mathcal{F}_{\tau_n}$ and has the same law as τ_1. It follows that the interarrival times are independent and exponential with parameter λ, as stated.

Hence the density f_{τ_n} of τ_n is

$$f_{\tau_n}(t) = \begin{cases} \lambda e^{-\lambda t} \dfrac{(\lambda t)^{n-1}}{(n-1)!} & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases}$$

i.e. τ_n has a gamma distribution[1] with parameters λ, n. (Prove this by induction: $f_{\tau_{n+1}}(t) = -\frac{d}{dt}\mathbb{P}(\tau_{n+1} > t) = -\frac{d}{dt}\int_{-\infty}^{\infty} \mathbb{P}(\tau_{n+1} - \tau_n \geq t - u | \tau_n = u) f_{\tau_n}(u)\, du = \ldots$)
Thus

$$\begin{aligned} \mathbb{P}(\tau_n \leq t) &= \int_0^t f_{\tau_n}(u)\, du \\ &= e^{-\lambda t}\frac{(\lambda t)^n}{n!} + \int_0^t \lambda e^{-\lambda u}\frac{(\lambda u)^n}{n!}\, du \qquad \text{integrating by parts} \\ &= e^{-\lambda t}\frac{(\lambda t)^n}{n!} + \mathbb{P}(\tau_{n+1} \leq t) \end{aligned}$$

and hence

$$\mathbb{P}(N_t = n) = \mathbb{P}(\tau_n \leq t) - \mathbb{P}(\tau_{n+1} \leq t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$

as required.

⊣

Remarks 4.4 (a) The above proof also indicates how to construct a Poisson process.: Let $\{\sigma_n : n \in \mathbb{N}\}$ be a family of independent exponential random variables with parameter λ. Let $\tau_n = \sum_{k=1}^n \sigma_k$. Then

$$N_t = \sum_{n \geq 1} I_{\{\tau_n \leq t\}}$$

defines a Poisson process.

(b) Note that, in that case,

$$\tau_n = \inf\{t \geq 0 : N_t \geq n\}$$

(c) The quantity $\lambda > 0$ is called the *arrival rate* or *intensity* of the Poisson process. It is easy too verify that

$$\mathbb{E}[N_t] = \lambda t = \mathrm{Var}(N_t)$$

Thus an intensity of λ corresponds to an expected λ arrivals per unit time.

(d) We do not need the all the axioms for a Lévy process to characterize a Poisson process: Any counting process with stationary independent increments is automatically continuous in probability, and thus a Lévy process. Indeed, it is shown in the proof above that $\mathbb{P}(N_t = 0) = e^{-\lambda t}$ for some $\lambda > 0$. It follows $\lim_{s \to t} \mathbb{P}(|N_s - N_t| > \varepsilon) = \lim_{s \to t} \mathbb{P}(|N_{t-s}| > \varepsilon) = \lim_{h \to 0} \mathbb{P}(N_h > \varepsilon) = \lim_{h \to 0}(1 - e^{-\lambda h}) = 0$.

□

[1] The Gamma distribution with parameters c, α has density $\gamma_{c,\alpha}(x) = \frac{\alpha^c}{\Gamma(c)} x^{c-1} e^{-\alpha x} I_{\mathbb{R}^+}(x)$ where $\Gamma(c) = \int_0^\infty t^{c-1} e^{-t}\, dt$

Next, we introduce *compound Poisson* processes. Let S_n be a random walk, i.e. $S_n = \sum_{k=1}^n Y_k$, where $Y_1, Y_2, \dots$ are i.i.d. random variables. Now effect a time change, so that the time between jumps is exponentially distributed: Let $\sigma_1, \sigma_2, \dots$ be independent exponential variables, also independent of $Y_1, Y_2, \dots$, let $\tau_n = \sum_{k=1}^n \sigma_k$, and let $N_t = \sum_n I_{\{\tau_n \le t\}}$ be the associated Poisson process. Define

$$X_t := S_{N_t} = \sum_{n=1}^{N_t} Y_n = \sum_{n \ge 1} Y_n I_{\{\tau_n \le t\}}$$

The process X_t is called a compound Poisson process. More precisely:

Definition 4.5 Suppose that $\{Y_n : n \in \mathbb{N}\}$ are i.i.d variables (with values in $\mathbb{R}^d$) with distribution μ and that N_t is a Poisson process with intensity λ (which is the counting process for the sequence of random times $(\tau_n : n \ge 1)$). Suppose further that $\{Y_n : n \in \mathbb{N}\}$ is independent of $\{N_t : t \ge 0\}$. The process

$$X_t := \sum_{n=1}^{N_t} Y_n = \sum_{n \ge 1} Y_n I_{\{N_t \ge n\}} = \sum_n Y_n I_{\{\tau_n \le t\}}$$

is called a *compound Poisson* process with intensity λ and jump distribution μ.

□

Here, it is understood that $\sum_{n=1}^0 Y_n := 0$. Note that a Poisson process is a one–dimensional compound Poisson process with jump distribution δ_1, the point mass with support $\{1\}$.

If $X_t = \sum_{k=1}^{N_t} Y_k$ is a compound Poisson process, with intensity λ and jump distribution μ, then the characteristic function of X_t is given by

$$\begin{aligned}
\mathbb{E}[e^{i\langle u, X_t \rangle}] &= \sum_{n=0}^{\infty} \mathbb{E}[e^{i\langle u, \sum_{k=1}^n Y_k \rangle} | N_t = n] \mathbb{P}(N_t = n) \\
&= \sum_{n=0}^{\infty} \left(\mathbb{E}[e^{i\langle u, Y \rangle}] \right)^n e^{-\lambda t} \frac{(\lambda t)^n}{n!} \\
&= e^{-\lambda t} \sum_{n=0}^{\infty} \frac{\hat{(}\mu(u)\lambda t)^n}{n!} \\
&= e^{\lambda t(\hat{\mu}(u) - 1)}
\end{aligned}$$

For that reason, we define:

Definition 4.6 A probability distribution σ on $\mathbb{R}^d$ is said to be *compound Poisson* if, for some $c > 0$ and some probability distribution μ on $\mathbb{R}^d$, we have

$$\hat{\sigma}(u) = e^{c(\hat{\mu}(u) - 1)}$$

□

Thus a Poisson distribution is a compound Poisson distribution with $c = 1$ and $\mu = \delta_1$.

Recall that a right–continuous function $f : \mathbb{R}_+ \to \mathbb{R}^d$ is said to be *piecewise constant* if there exist $0 = t_0 < t_1 < t_2 < \dots$ and constants $a_n \in \mathbb{R}^d$ such that $f = \sum_n a_n I_{[t_n, t_{n+1})}$, where each bounded interval in $\mathbb{R}^+$ contains at most finitely many of the t_n. The class of compound Poisson processes is exactly the class of Lévy processes with a.s. piecewise constant sample paths.

Theorem 4.7 *A process X_t is a compound Poisson process if and only if it is a Lévy process with a.s. piecewise constant sample paths.*

Proof: ($\Rightarrow$): Suppose that $X_t = \sum_{n=1}^{N_t} Y_n$ is a compound Poisson process, where N_t is a Poisson process with intensity λ, and the Y_n are mutually independent with common distribution μ, and also independent of $(N_t)_{t\geq 0}$. Let $S_n = \sum_{k=1}^{n} Y_k$ be the associated random walk, so that $X_t = S_{N_t}$. Since X_t jumps only when N_t does, X_t is càdlàg, because N_t is. Since N_t is piecewise constant (because $\mathbb{P}(N_t < \infty) = 1$), so is X_t.

Suppose now that $B_0, B_1 \dots \in \mathcal{B}(\mathbb{R}^d)$ and that $0 \leq t_0 < t_1 < t_2 < \dots$. Then we have, by stationary independent increments for the Poisson process and the random walk, that

$$
\begin{aligned}
&\mathbb{P}(X_{t_0} \in B_0, X_{t_1} - X_{t_0} \in B_1, \dots X_{t_k} - X_{t_{k-1}} \in B_k) \\
&= \sum_{n_0, n_1, \dots n_k \in \mathbb{N}} \mathbb{P}\Big(N_{t_0} = n_0, N_{t_1} - N_{t_0} = n_1, \dots, N_{t_k} - N_{t_{k-1}} = n_k, \\
&\qquad\qquad S_{n_0} \in B_0, S_{n_0+n_1} - S_{n_0} \in B_1, \dots, S_{n_0+\dots+n_k} - S_{n_0+\dots+n_{k-1}} \in B_k\Big) \\
&= \sum_{n_0, n_1, \dots n_k \in \mathbb{N}} \mathbb{P}(N_{t_0} = n_0)\mathbb{P}(N_{t_1 - t_0} = n_1) \dots \mathbb{P}(N_{t_k - t_{k-1}} = n_k) \cdot \\
&\qquad\qquad \mathbb{P}(S_{n_0} \in B_0) \cdot \mathbb{P}(S_{n_1} \in B_1) \dots \mathbb{P}(S_{n_k} \in B_k) \\
&= \mathbb{P}(X_{t_0} \in B_0) \cdot \mathbb{P}(X_{t_1 - t_0} \in B_1) \dots \mathbb{P}(X_{t_k - t_{k-1}} \in B_k)
\end{aligned}
$$

In particular,

$$
\mathbb{P}(X_{t_1} - X_{t_0} \in B) = \mathbb{P}(X_{t_0} \in \mathbb{R}^d, X_{t_1} - X_{t_0} \in B) = \mathbb{P}(X_{t_0} \in \mathbb{R}^d)\mathbb{P}(X_{t_1 - t_0} \in B)
$$

which shows that $X_{t_1} - X_{t_0}$ and $X_{t_1 - t_0}$ are identical in law, and thus that X_t has stationary increments. It follows that

$$
\begin{aligned}
&\mathbb{P}(X_{t_0} \in B_0, X_{t_1} - X_{t_0} \in B_1, \dots X_{t_k} - X_{t_{k-1}} \in B_k) \\
&= \mathbb{P}(X_{t_0} \in B_0) \cdot \mathbb{P}(X_{t_1} - X_{t_0} \in B_1) \dots \mathbb{P}(X_{t_k} - X_{t_{k-1}} \in B_k)
\end{aligned}
$$

and thus that X_t has independent increments.

($\Leftarrow$): Now suppose that X_t is a Lévy process, with piecewise continuous sample paths. Define a sequence of stopping times $(\tau_n)_{n\geq 0}$ by

$$
\tau_0 = 0 \qquad \tau_{n+1} = \inf\{t > \tau_n : X_{t-} \neq X_t\}
$$

Since X_t has a.s. piecewise constant sample paths, the τ_n are a.s. finite and form a strictly increasing sequence of stopping times. Let N_t be the associated counting process, so that

$$
N_t = \text{no. of jumps of } X \text{ in interval } (0, t]
$$

Then $N_t - N_s =$ no. of jumps of X in interval $(s, t]$. Because X is a Lévy process this is identical in law to $N_{t-s} =$ the no. of jumps in interval $(0, t-s]$, and independent of $\sigma(X_u : u \leq s)$. In summary, N_t is a Lévy process, because X_t is a Lévy process. Since N_t is also a counting process, it must be a Poisson process, by Thm. 4.3.

Now define, $Y_n = X_{\tau_n} - X_{\tau_n -}$ (where $n \geq 1$). Then Y_n is the first jump of the process $X_t^n := X_{\tau_{n-1}+t} - X_{\tau_{n-1}}$, which is a Lévy process identical in law to X, and independent of $\mathcal{F}_{\tau_{n-1}}$, by the strong Markov property. It follows that the Y_n are i.i.d. The jump sizes must also be independent of the no. of jumps (or else X_t could not have independent increments), and thus the Y_n are independent of N_t.

$\dashv$

Remarks 4.8 Since every càdlàg function can be approximated by piecewise constant functions, it is not hard to believe that every Lévy process can be approximated by compound Poisson processes. Much more about this later.

$\square$

5 Lévy Processes without Jumps

The aim of this section is to show that a continuous Lévy process in $\mathbb{R}^d$ is none other than an arithmetic Brownian motion in a loose sense: The components are arithmetic Brownian motions that need not be independent, but their joint distribution is multivariate Gaussian.

Our aim is to use the Lévy characterization of Brownian motion: A one–dimensional continuous martingale M vanishing at $t = 0$ is a standard Brownian motion iff $[M]_t = t$, where $[M]_t$ is the quadratic variation of M. Cf. Revuz and Yor[**?**] Thm IV.3.6 for a proof. Also see Appendix D for a development which does not rely on Lévy's characterization theorem.

In order to apply this result, we need to know that the first and second moments of a Lévy process with continuous sample paths exist. The following result gives us more than we need:

Proposition 5.1 *A Lévy process whose jumps are a.s. bounded has moments of all orders.*

Proof: Let X_t be a Lévy process, and suppose that K is a bound for the jumps, i.e. that $\mathbb{P}(\exists t[|X_t - X_{t-}| \geq K]) = 0$. Define stopping times $T_n, n \in \mathbb{N}$ inductively by

$$T_1 = \inf\{t : |X_t| \geq K\}, \quad T_{n+1} = \inf\{t : t > T_n, |X_t - X_{T_n}| \geq K\}$$

By right–continuity of the process X_t, the sequence $(T_n)_n$ is a.s. strictly increasing. Since $|X_{T_n} - X_{T_n -}| \leq K$ for all n, we see by induction that $\sup_{s \leq T_n} |X_s| \leq 2nK$: Indeed, $\sup_{s < T_1} |X_s| \leq K$, and hence $\sup_{s \leq T_1} |X_s| \leq K + K$, using $|X_{T_1} - X_{T_1 -}| \leq K$. The result for $n > 1$ follows by similar reasoning.

By the strong Markov property, $T_n - T_{n-1}$ is independent of $\mathcal{F}_{T_{n-1}}$, and have the same law as T_1. Thus

$$\mathbb{E}[e^{-T_n}] = \mathbb{E}[e^{-T_1}]^n = a^n \qquad \text{where } a < 1$$

It follows that

$$\mathbb{P}(|X_t| \geq 2nK) \leq \mathbb{P}(T_n < t) \leq e^t a^n$$

(because $\mathbb{E}[e^{-Z}] \geq \mathbb{E}[e^{-Z}; Z < t] \geq e^{-t}\mathbb{P}(Z < t)$.) Thus

$$\begin{aligned}\mathbb{E}[|X_t|^m] &= \sum_n \mathbb{E}[|X_t|^m; 2nK < |X_t| \leq 2(n+1)K] \\ &= e^t (2K)^m \sum_n (n+1)^m \alpha^n < \infty\end{aligned}$$

by the ratio test.

$\dashv$

Since for distributions μ we have that $\int |x|^m \, d\mu < \infty$ implies that $\hat{\mu} \in C^m$ (cf. Thm. 2.3(h)), we see:

Corollary 5.2 *If X_t is a Lévy process whose jumps are a.s. bounded, then the map*

$$u \mapsto \mathbb{E}[e^{i\langle u, X_t \rangle}]$$

is C^∞.

$\square$

Suppose that X_t is a d–dimensional Lévy process with continuous sample paths. Then it has moments of all orders, by Propn. 5.1. In particular, the process $X_t - \mathbb{E}[X_t]$ is a continuous martingale centered at 0. We now show that any centered continuous Lévy process is a Brownian motion in the loose sense: Components need not be independent, but are multi–variate Gaussian.

So let $X_t = (X_t^{(1)}, \dots, X_t^{(d)})$ be a centered Lévy process with continuous sample paths. Let A be the non–negative definite symmetric $d \times d$–matrix defined by

$$A_{ij} = \mathbb{E}[X_1^{(i)} X_1^{(j)}]$$

where $X_t = (X_t^{(1)}, \dots, X_t^{(d)})$. We claim that the covariation process of $X_t^{(i)}$ and $X_t^{(j)}$ is given by

$$[X^{(i)}, X^{(j)}]_t = A_{ij} t$$

Recall that $\mathbb{E}[e^{i\langle u, X_t \rangle}] = e^{t\psi(u)}$ for some $\psi : \mathbb{R}^d \to \mathbb{C}$ with $\psi(0) = 0$. Since X_t has moments of all orders, the function ψ is $\mathcal{C}^\infty$, by Thm. 2.3(h). Moreover, since $\mathbb{E}[X_t^{(i)}] = 0$, we have $\frac{\partial}{\partial u^{(i)}}\Big|_{u=0} e^{t\psi(u)} = 0$, and thus that $\frac{\partial}{\partial u^{(i)}}\Big|_{u=0} \psi(u) = 0$ also. It now follows easily that

$$\mathbb{E}[X_t^{(i)} X_t^{(j)}] = -\frac{\partial^2}{\partial u^{(i)} \partial u^{(j)}}\Big|_{u=0} e^{t\psi(u)} = t \, \frac{\partial^2}{\partial u^{(i)} \partial u^{(j)}}\Big|_{u=0} \psi(u) = t \, \mathbb{E}[X_1^{(i)} X_1^{(j)}]$$

i.e. that

$$\mathbb{E}[X_t^{(i)} X_t^{(j)}] = A_{ij} t$$

To show that $[X^{(i)}, X^{(j)}]_t = A_{ij} t$, it suffices to show that $X_t^{(i)} X_t^{(j)} - A_{ij} t$ is a martingale. By the fact that increments are independent with mean zero, we have

$$\begin{aligned}&\mathbb{E}[X_t^{(i)} X_t^{(j)} - X_s^{(i)} X_s^{(j)} | \mathcal{F}_s] \\ &= \mathbb{E}[(X_t^{(i)} - X_s^{(i)})(X_t^{(j)} - X_s^{(j)}) + X_s^{(i)}(X_t^{(j)} - X_s^{(j)}) + X_s^{(j)}(X_t^{(i)} - X_s^{(i)}) | \mathcal{F}_s] \\ &= \mathbb{E}[(X_t^{(i)} - X_s^{(i)})(X_t^{(j)} - X_s^{(j)})]\end{aligned}$$

Taking expectations on both sides shows that

$$\mathbb{E}[X_t^{(i)}X_t^{(j)} - X_s^{(i)}X_s^{(j)}] = \mathbb{E}[(X_t^{(i)} - X_s^{(i)})(X_t^{(j)} - X_s^{(j)})]$$

It follows that

$$\mathbb{E}[X_t^{(i)}X_t^{(j)} - X_s^{(i)}X_s^{(j)}|\mathcal{F}_s] = \mathbb{E}[X_t^{(i)}X_t^{(j)} - X_s^{(i)}X_s^{(j)}] = A_{ij}(t-s)$$

and thus that $X_t^{(i)}X_t^{(j)} - A_{ij}t$ is a martingale.

Now, for $\lambda \in \mathbb{R}^d$, let $Z_t^\lambda = \langle \lambda, X_t \rangle$. Then Z_t^λ is a centered continuous one–dimensional martingale. Using the fact that the covariance process bracket operation is bilinear, we see that the quadratic variation of Z_t^λ is given by

$$[Z^\lambda]_t = \langle \lambda, A\lambda \rangle \; t$$

Hence, by Lévy's characterization, Z_t^λ is a Brownian motion with variance parameter $\langle \lambda, A\lambda \rangle$ (i.e. $Z_t^\lambda \sim N(0, \langle \lambda, A\lambda \rangle \; t)$). It now follows that

$$\mathbb{E}[e^{i\langle u, X_t \rangle}] = \varphi_{Z_t^u}(1) = e^{-\frac{1}{2}\langle u, Au \rangle \; t}$$

which proves that X_t is a d–dimensional Brownian motion with covariance matrix A.

Now if X is a Lévy process with continuous sample paths, then $X_t - \mathbb{E}[X_t]$ is centered, and thus a Brownian motion. If we define $\gamma = \mathbb{E}[X_1]$, then $\mathbb{E}[X_t] = \gamma t$, and so $\mathbb{E}[e^{i\langle u, X_t - \gamma t \rangle}] = e^{-\frac{1}{2}\langle u, Au \rangle \; t}$ for some symmetric non–negative definite matrix A. We have proved:

Theorem 5.3 *Suppose that X_t is a Lévy process with continuous sample paths. Then there exists $\gamma \in \mathbb{R}^d$ and a symmetric non–negative definite $d \times d$–matrix A such that*

$$\mathbb{E}[e^{i\langle u, X_t \rangle}] = e^{i\langle u, \gamma \rangle \; t - \frac{1}{2}\langle u, Au \rangle \; t}$$

Hence X_t is a d–dimensional arithmetic Brownian motion with drift γ and covariance matrix A.

□

6 Analysis of Jumps of a Lévy Process

6.1 A Useful Technical Result

Recall that if f is a function $f : [0, \infty) \to \mathbb{R}^d$, and if $0 \leq t_1 < t_2 < \infty$, then the variation $V(f; (t_1, t_2])$ of f over the interval $(t_1, t_2]$ is defined by

$$V(f; (t_1, t_2]) = \sup_\Delta \sum_{j=1}^n |f(s_j) - f(s_{j-1})|$$

The following result will play an important role in our analysis of the jumps of a Lévy process:

Proposition 6.1 *Suppose that M, N are centered càdlàg martingales, that M is square–integrable, and that N has square–integrable variation (i.e. $\mathbb{E}[V_t(N)^2] < \infty$, where $V_t(N)$ is the variation of N on $[0, t]$). Then*

$$\mathbb{E}M_t N_t = \mathbb{E}\sum_{s \leq t} \Delta M_s \cdot \Delta N_s$$

Proof: A rigorous proof is given in Appendix E. The basic idea, however, is simple: Note that if $\pi = \{0 = t_0 < t_1 < \cdots < t_N = t\}$ is a partition of $[0, t]$, then

$$\mathbb{E}M_tN_t = \mathbb{E}\left(\sum_{0\le k<N}(M_{t_{k+1}} - M_{t_k})\sum_{0\le k<N}(N_{t_{k+1}} - N_{t_k})\right) = \mathbb{E}\left[\sum_{0\le k<N}(M_{t_{k+1}} - M_{t_k})(N_{t_{k+1}} - N_{t_k})\right]$$

by repeated applications of the tower property. It is not hard to believe that

$$\sum_{0\le k<N}(M_{t_{k+1}} - M_{t_k})(N_{t_{k+1}} - N_{t_k}) \to \sum_{s\le t}\Delta M_s \cdot \Delta N_s \qquad \text{as mesh}(\pi) \to 0$$

which yields the result, provided we may apply the dominated convergence theorem. The technical conditions ensure that we may. Full details are in Appendix E.

$\dashv$

6.2 Counting Jumps

Let $X = (X^1, \ldots, X^d)$ be a d–dimensional Lévy process on a space $(\Omega, \mathcal{F}, \mathbb{P})$. Recall that the jump process is given by $\Delta X_t = X_t - X_{t-}$. This definition makes sense, because X is càdlàg. It is quite possible for $\sum_{s\le t}\Delta X^i_s$ to be infinite. However, again because X is càdlàg, on any bounded time interval there are only finitely many jumps whose amplitude exceeds a given (strictly positive) size. Put another way: Suppose that $B \in \mathcal{B}(\mathbb{R}^d)$, and that $0 \notin \bar{B}$. Then

$$\sum_{s\le t}\Delta X_s I_{\{\Delta X_s\in B\}}$$

has only finitely many non–zero terms. We say that a Borel set B is *bounded away from zero* if $0 \notin \bar{B}$. For such B, we can introduce a strictly increasing sequence of stopping times $(\tau^B_n)_{n\in\mathbb{N}}$ by

$$\tau^B_0 = 0 \qquad \tau^B_{n+1} = \inf\{t > \tau^B_n : \Delta X_t \in B\}$$

Thus the $(\tau^B_n)_{n\ge1}$ enumerate the times of jumps in B. Let $N_t(B)$ be the associated counting process:

$$N_t(B) = \sum_{n=1}^{\infty} I_{\{\tau^B_n\le t\}} = \sum_{0<s\le t} I_B(\Delta X_s)$$

Let $\nu(B)$ be the parameter of $N_t(B)$, i.e.

$$\nu(B) = \mathbb{E}[N_1(B)]$$

is the expected no. of jumps of $N_t(B)$ per unit time. Then $\nu(B)$ is also the expected no. of jumps of X in B per unit time. Note that also that $\nu(B) < \infty$ when B is bounded away from zero. It is easy to see that ν is a measure on $\mathcal{B}(\mathbb{R}^d\backslash\{0\})$: For fixed $\omega \in \Omega$, the map

$$N_t(\cdot)(\omega) : \mathcal{B}(\mathbb{R}^d\backslash\{0\}) \to \mathbb{N} \cup \{\infty\} : B \mapsto N_t(B)(\omega)$$

is a counting measure, counting the no. of jumps of X in B by time t. Since $\nu(B) = \mathbb{E}[N_1(B)]$, the monotone convergence theorem implies that ν is a measure also.

Proposition 6.2 *Let X be a Lévy process in $\mathbb{R}^d$. For each $B \in \mathcal{B}(\mathbb{R}^d)\backslash\{0\}$, let $N_t(B) = \sum_{s\leq t} I_B(\Delta X_s)$. Then*

(a) $N_t(B)$ is a Poisson process with intensity $\nu(B)$.

(b) If $B_1, \ldots B_m \in \mathcal{B}(\mathbb{R}^d\backslash\{0\})$ are disjoint, then $N_t(B_1), N_t(B_2), \ldots, N_t(B_m)$ are independent.

Proof:

We verify that $N_t(B)$ is a Poisson process: It is easy to see that $N_t(B)$ is càdlàg, because N jumps only when X jumps, and X is càdlàg. Furthermore, $N_t(B) - N_s(B)$ is the number of jumps in the interval $(s,t]$, which is clearly $\sigma(X_u - X_v : s \leq u < v)$–measurable, and therefore independent of $\mathcal{F}_s$, because X has independent increments. Thus $N_t(B)$ has independent increments. Because the no. of jumps in $(s,t]$ and the no. jumps in $(0, t-s]$ are identically distributed (because the processes X_u and $X_{u+s} - X_s$ are identical in law), we see that $N_t(B) - N_s(B)$ has the same distribution as $N_{t-s}(B)$. Thus $N_t(B)$ is a Lévy process which is also a counting process, and hence a Poisson process.

Furthermore, if λ_B is the intensity of the Poisson process $N_t(B)$, then, in particular $\mathbb{E}[N_1(B)] = \lambda_B \cdot 1$, i.e. $\lambda_B = \nu(B)$. This proves (a).

Now suppose that $B, C \in \mathcal{B}(\mathbb{R}^d)$ are disjoint and bounded away from zero. Let

$$M_B^u(t) = \frac{e^{iuN_t(B)}}{\mathbb{E}[e^{iuN_t(B)}]} - 1 \qquad M_C^v(t) = \frac{e^{iuN_t(C)}}{\mathbb{E}[e^{iuN_t(C)}]} - 1$$

Then M_B^u, M_C^v are centered càdlàg martingales. Note that $M_B^u(t)$ jumps only when $N_t(B)$ jumps. Since $N_t(B)$ and $N_t(C)$ cannot jump at the same time (because $B \cap C = \emptyset$), we have $\Delta M_B^u(s)\ \Delta M_C^v(s) = 0$. By Propn. 6.1, we see that $\mathbb{E}[M_B^u(t) M_C^v(t)] = 0$. It follows immediately that

$$\mathbb{E}[e^{iuN_t(B) + ivN_t(C)}] = \mathbb{E}[e^{iuN_t(B)}]\ \mathbb{E}[e^{ivN_t(C)}]$$

which proves that $N_t(B), N_t(C)$ are independent, by Thm. 2.3(Kac's Theorem).

Also, $M_B^u(t) M_C^v(t)$ is then clearly a centered martingale. It is now easy to show (by induction) that $N_t(B_1), \ldots, N_t(B_m)$ are independent whenever $B_1, \ldots, B_m$ are disjoint and bounded away form zero.

It remains to prove the result for general Borel subsets of $\mathbb{R}^d\backslash\{0\}$. If $B \in \mathcal{B}(\mathbb{R}^d\backslash\{0\})$, then we can write $B = \bigcup_n D_B^n$ as a union of Borel sets that are bounded away from zero, where $D_B^n = \{x \in B : |x| \geq \frac{1}{n}\}$. Then $N_t(B) = \lim_n N_t(D_B^n)$ a.s. , where each $N_t(D_B^n)$ is Poisson with mean $\nu(D_B^n)$. Since a.s. convergence implies convergence in distribution, we see, by comparing characteristic functions, then $N_t(B)$ is Poisson with mean $\lim_n \nu(D_B^n) = \nu(B)$. Finally, if $B_1, \ldots, B_m \in \mathcal{B}(\mathbb{R}^d\backslash\{0\})$ are disjoint, then $N_t(D_{B_1}^{n_1}), \ldots, N_t(D_{B_m}^{n_m})$ are independent, for any $n_1, \ldots, n_m \in \mathbb{N}$. Since $N_t(B) \in \sigma(N_t(D_B^n) : n \in \mathbb{N})$, it follows that $N_t(B_1), \ldots, N_t(B_m)$ are independent also.

$\dashv$

Proposition 6.3 *ν is a σ–finite measure on $\mathbb{R}^d\backslash\{0\}$.*

Proof: This follows from the fact that $\nu(B) < \infty$ for any B bounded away from zero.

$\dashv$

We call ν the *Lévy measure* of the process X. Note that it is defined on the σ–algebra $\mathcal{B}(\mathbb{R}^d\backslash\{0\})$. It is convenient to extend ν to all of $\mathcal{B}(\mathbb{R}^d)$ by putting

$$\nu(\{0\}) = 0$$

6.3 Poisson Random Measures

Definition 6.4 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $(E, \mathcal{B})$ be a measurable space.

(a) A map $M : \mathcal{B} \times \Omega \to \mathbb{R}$ is called a *random measure* on $(E, \mathcal{B})$ iff

  (i) For each $B \in \mathcal{B}$, the map $\omega \mapsto M(B, \omega)$ is a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$.

  (ii) For almost every $\omega \in \Omega$, the map $B \mapsto M(B, \omega)$ is a measure on $(E, \mathcal{B})$.

  (iii) There exists a partition $B_1, B_2, \dots \in \mathcal{B}$ of E such that $M(B_k) < \infty$ a.s. for all k. (This is a σ–finiteness condition.)

(b) A random measure M on $(E, \mathcal{B})$ is said to have *independent increments* iff $M(B_1), \dots, M(B_n)$ are independent random variables whenever $B_1, \dots, B_n$ are mutually disjoint members of $\mathcal{B}$.

(c) A random measure M on $(E, \mathcal{B})$ is called a *point process* iff M is a $\bar{\mathbb{Z}}_+$–valued (where $\bar{\mathbb{Z}}_+ = \{0, 1, 2, \dots, \infty\}$).

(d) Let μ be a σ–finite measure on $(E, \mathcal{B})$. A *Poisson random measure* with *intensity measure* μ is a point process M with independent increments such that for every $B \in \mathcal{B}$, $M(B)$ is a Poisson random variable with mean $\mu(B)$, i.e.

$$\mathbb{P}(M(B) = k) = e^{-\mu(B)}\frac{\mu(B)^k}{k!} \qquad \text{for all } k \in \bar{\mathbb{Z}}_+$$

Here, we use the conventions that if $\mu(B) = 0$, then $\mathbb{P}(M(B) = 0) = 1$, and if $\mu(B) = \infty$, then $\mathbb{P}(M(B) = \infty) = 1$.

$\square$

Poisson random measures have already been encountered before:

Proposition 6.5 *Let X be a Lévy process in $\mathbb{R}^d$. For each $B \in \mathcal{B}(\mathbb{R}^d)\backslash\{0\}$, let $N_t(B) = \sum_{s\le t} I_B(\Delta X_s)$ be the number of jumps in B by time t. Define the measure ν on $\mathcal{B}(\mathbb{R}^d\backslash\{0\})$ by $\nu(B) = \mathbb{E}[N_1(B)]$. Then N_t is a Poisson random measure with intensity measure $t\nu$.*

Proof: Propn. 6.2.

$\dashv$

Next, we deal with the matter of existence of Poisson random measures:

Theorem 6.6 *Suppose that $(E, \mathcal{B}, \mu)$ is σ–finite measure space. Then there exists, on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a Poisson random measure M on $(E, \mathcal{B})$ with intensity measure μ.*

$\square$

The proof may be found in appendix F.

Suppose that $X \subseteq \mathbb{R}^d$. By a *rectangle* in X, we mean a set of the form $(a_1, b_1] \times \cdots \times (a_d, b_d] \cap X$. The following lemma states that a random measure which behaves like a Poisson random measure on rectangles is, in fact, a random measure.

Lemma 6.7 *Suppose that $N : \mathcal{B}(X) \times \Omega \to \bar{\mathbb{R}}_+$ is a random measure on a Borel subset X of $\mathbb{R}^d$, and suppose that the intensity μ of N (given by $\mu(B) = \mathbb{E}[N(B)]$ for $B \in \mathcal{B}(\mathbb{R}^d)$) is σ–finite. Suppose that*

(i) $N(B)$ is a Poisson random variable for each rectangle *$B \in \mathcal{B}(\mathbb{R}^d)$;*

(ii) If $B_1, B_2, \ldots, B_n$ are disjoint rectangles, then $N(B_1), N(B_2), \ldots, N(B_n)$ are independent.

Then N is a Poisson random measure with intensity μ.

□

6.4 Integration with Respect to a Poisson Random Measure

The following important result is taken from Sato[**?**].

Proposition 6.8 *Let $(E, \mathcal{B}, \mu)$ be a finite measure space, and let $N : \mathcal{B} \times \Omega \to \bar{\mathbb{N}}$ be a Poisson random measure with intensity measure μ. Let $f : (E, \mathcal{B}) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ be measurable.*

(a) Then

$$Z(\omega) = \int_E f(x)\; N(dx, \omega)$$

is a random variable (with values in $\mathbb{R}^d$) with characteristic function

$$\begin{aligned} \mathbb{E}[e^{i\langle u, Z\rangle}] &= \exp(\left(\int_E e^{i\langle u, f(x)\rangle} - 1\; \mu(dx)\right) \\ &= \exp\left(\int_{\mathbb{R}^d} e^{i\langle u, x\rangle} - 1\; (\mu f^{-1})(dx)\right) \end{aligned}$$

i.e. Z is a **compound Poisson RV** *with jump distribution $\frac{\mu f^{-1}}{\mu f^{-1}(\mathbb{R}^d)}$ and intensity $\mu f^{-1}(\mathbb{R}^d)$.*

(b) If each $\int_E |f^{(j)}(x)|\; \mu(dx) < \infty$ (where $f^{(j)}$ is the j^{th} coordinate of f), then

$$\mathbb{E}[Z] = \int_E f(x)\; \mu(dx)$$

(c) If $\int_E |f(x)|^2\; \mu(dx) < \infty$, then

$$\mathbb{E}\left[\left|Z - \mathbb{E}[Z]\right|^2\right] = \int_E |f(x)|^2\; \mu(dx)$$

(d) If $B_1, \ldots, B_m \in \mathcal{B}$ are disjoint, then

$$Z_k = \int_{B_k} f(x)\, N(dx, \omega)$$

are independent random variables.

Proof: Note that $Z < \infty$ a.s. because $N(E)$ is finite a.s. (and hence supported on a finite number of points). We can approximate Z by simple random variables as follows. Given a point $p = (p_1, \ldots, p_d) \in \mathbb{Z}^d$, let C_p^n be the rectangle in $\mathbb{R}^d$ consisting of all points $y = (y_1, \ldots, y_d)$ satisfying $2^{-n}(p_j - 1) < y_j \leq 2^{-n} p_j$. Let y_p^n be a point in C_p^n, and define $f_n : E \to \mathbb{R}^d$ by

$$f_n(x) = \sum_{p \in \mathbb{Z}^d} y_p^n I_{C_p^n}(f(x))$$

Also define

$$Z_n(\omega) = \int_E f_n(x)\, N(dx, \omega)$$

Then $|f(x) - f_n(x)| \leq 2^{-n}\sqrt{d}$, and so $|Z_n(\omega) - Z(\omega)| \leq 2^{-n}\sqrt{d} N(E, \omega)$. It follows that $|Z_n(\omega) - Z(\omega)| \to 0$ a.s. as $n \to \infty$. Now since $Z_n(\omega) = \sum_{p \in \mathbb{Z}^d} y_p^n\, N(f^{-1}(C_p^n), \omega)$, we have

$$\begin{aligned}
\mathbb{E}[e^{i\langle u, Z_n\rangle}] &= \prod_{p \in \mathbb{Z}^d} \mathbb{E}[e^{i\langle u, y_p^n\rangle N(f^{-1}(C_p^n))}] \\
&= \prod_{p \in \mathbb{Z}^d} \exp\left(\mu(f^{-1}(C_p^n))[e^{i\langle u, y_p^n\rangle} - 1]\right) \\
&= \exp\left(\int_E e^{i\langle u, f_n(x)\rangle} - 1\ \mu(dx)\right)
\end{aligned}$$

using the definition of Poisson random measures. A.s. convergence implies weak convergence, so

$$\mathbb{E}[e^{i\langle u, Z\rangle}] = \lim_n \mathbb{E}[e^{i\langle u, Z_n\rangle}] = \exp\left(\int_E e^{i\langle u, f(x)\rangle} - 1\ \mu(dx)\right)$$

Let $u^{(j)}, Z^{(j)}, f^{(j)}$ denote the j^{th} coordinates of u, Z, f. Assuming $\int |f^{(j)}(x)|\, \mu(dx) < \infty$, we see that

$$\frac{\partial}{\partial u^{(j)}} \int e^{i\langle u, f(x)\rangle} - 1\ \mu(dx) = i \int f^{(j)}(x) e^{i\langle u, f(x)\rangle}\ \mu(dx)$$

Hence

$$\mathbb{E}[Z^{(j)}] = i^{-1} \frac{\partial}{\partial u^{(j)}} \exp\left(\int e^{i\langle u, f(x)\rangle} - 1\ \mu(dx)\right)\Big|_{u=0} = \int f^{(j)}(x)\ \mu(dx)$$

Similarly, when $\int |f(x)|^2\, \mu(dx) < \infty$, then

$$\frac{\partial^2}{\partial (u^{(j)})^2} \int e^{i\langle u, f(x)\rangle} - 1\ \mu(dx) = i^2 \int f^{(j)}(x)^2 e^{i\langle u, f(x)\rangle}\ \mu(dx)$$

and thus

$$\begin{aligned}
\mathbb{E}[(Z^{(j)})^2] &= i^{-2} \frac{\partial^2}{\partial (u^{(j)})^2} \exp\left(\int e^{i\langle u, f(x)\rangle} - 1\ \mu(dx)\right)\Big|_{u=0} \\
&= \left(\int f^{(j)}(x)\ \mu(dx)\right)^2 + \int f^{(j)}(x)^2\ \mu(dx) \\
&= \mathbb{E}[Z^{(j)}]^2 + \int f^{(j)}(x)^2\ \mu(dx)
\end{aligned}$$

Finally, if $B_1, \dots, B_m \in \mathcal{B}$ are disjoint, then define

$$Z_{k,n} = \int_{B_k} f_n(x)\; N(dx, \omega) \qquad \text{for } k = 1, \dots, m$$

where the f_n are defined above. Then $Z_{k,n}(\omega) = \sum_{p \in \mathbb{Z}^d} y_p^n N(B_k \cap C_p^n)$. Since the $N(B_k \cap C_p^n)$ are independent for $k = 1, \dots, m$ and $p \in \mathbb{Z}^d$, we see that $Z_{1,n}, Z_{2,n}, \dots, Z_{m,n}$ are independent (for all n). Now $Z_{k,n} \to \int_{B_k} f(x)\; N(dx)$ a.s. as $n \to \infty$, and thus the $\int_{B_k} f(x)\; N(dx)$ are independent also.

$\dashv$

6.5 Jump Measure of a Lévy Process

Let $H = (0, \infty) \times \mathbb{R}^d \backslash \{0\}$. Every Lévy process X has a Poisson random measure J_X on $(H, \mathcal{B}(H))$ associated with it: Define

$$J_X(A, \omega) = \#\{t : (t, \Delta X_t) \in A\} \qquad \text{for } A \in \mathcal{B}(H), \omega \in \Omega$$

This is just a counting measure, and thus certainly a measure. In particular, if $A = (0, t] \times B$, where $B \in \mathcal{B}(\mathbb{R}^d)$ is bounded away from zero, then

$$J_X((0, t] \times B]) = \text{ no. of jumps of } X \text{ in } B \text{ by time } t = N_t(B)$$

Proposition 6.9 *Suppose that X is a Lévy process with Lévy measure ν. Then J_X is a Poisson random measure on $(H, \mathcal{B}(H))$ with intensity measure $\lambda \times \nu$.*

Proof: Let $\mu = \lambda \times \nu$. If $A = (s, t] \times B$, where $0 \le s < t$ and $B \in \mathcal{B}(\mathbb{R}^d)$ is bounded away from zero, then $J_X(A) = N_t(B) - N_s(B)$. Since $(N_t(B))_{t \ge 0}$ is a Poisson process with parameter $\nu(B)$, it follows that

$$\mathbb{P}(J_X(A) = k) = e^{-\nu(B)(t-s)} \frac{[\nu(B)(t-s)]^k}{k!} = e^{-\mu(A)} \frac{\mu(A)^k}{k!}$$

In particular, J_X is a random measure with independent increments with the property that $J_X(A)$ is Poisson with mean $\mu(A)$ for every rectangle $A \subseteq H$. Moreover, it is not hard to verify that $J_X(A_1), \dots, J_X(A_m)$ are independent if $A_1, \dots, A_m$ are disjoint rectangles. By manipulating the sets, we need only consider two cases:
Case 1: $A_j = (s, t] \times B_j$, where $B_1, \dots, B_m$ are mutually disjoint. In that case $J_X(A_j) = N_t(B_j) - N_s(B_j)$, and these are independent by Propn. 6.2(b).
Case 2: $A_j = (s_j, t_j] \times B_j$, where $(s_1, t_1], \dots, (s_m, t_m]$ are mutually disjoint intervals. In that case, $J_X(A_1), \dots, J_X(A_m)$ are independent by Propn. 6.2(a), because the $N_t(B_j)$ have independent increments.

Thus, lemma 6.7, J_X is a Poisson random measure with intensity μ.

$\dashv$

7 The Lévy–Ito Decomposition Theorem

We give present here a heuristic "proof" of the Lévy–Ito Decomposition Theorem in order to understand why is is true. Suppose that X_t is a one–dimensional Lévy process. Firstly, we ignore all jumps whose amplitude is less than a given lower bound ε, i.e. we assume that X only has jumps of size $\geq \varepsilon$. Then there is a strictly increasing sequence of stopping times $0 \leq \tau_1 < \tau_2 < \dots$ which enumerate the jump times of X. Let $Y_n = \Delta X_{\tau_n}$ denote the n^{th} jump of X, and let $N_t = \sum_{n=1}^{\infty} I_{\{\tau_n \leq t\}}$ be the counting process for the τ_n. Define $X_t^d = \sum_{n=1}^{N_t} Y_n$ to be the jump part of X. Then we can write $X_t = X_t^c + X_t^d$, where X^c is continuous.

Now because X is a Lévy process, so is N_t. It follows that N_t is a Poisson process. Again, because X has independent identically distributed increments, the Y_n are independent and identically distributed, and hence X_t^d is a compound Poisson process. Suppose that N_t has parameter λ, and that the distribution of the Y_n is σ. Suppose that $\bar{\sigma} := \mathbb{E}[Y]$ is the common mean of the Y_n. Then $\mathbb{E}[X_t^d] = \bar{\sigma}\lambda t$, and $J_t = X_t^d - \bar{\sigma}\lambda t$ is a martingale.

It is not too difficult to believe that X_t^c is a Lévy process. Let $\mu := \mathbb{E}[X_1^c]$. By the properties of Lévy processes, $\mathbb{E}[X_t^c] = \mu t$. Now define $M_t = X_t^c - \mu t$. Then M_t is a centered continuous Lévy process, and thus a continuous martingale. The fact that M_t is a Lévy process implies that the quadratic variation of M_t is proportional to t. For example $[M]_2 = \lim_{||P|| \to 0} \sum_k (M_{t_{k+1}} - M_{t_k})^2$ should be identical in distribution to the sum of two independent copies of $[M]_1$, i.e. the quadratic variation of M over $[0, 2]$ should be twice the quadratic variation of M over $[0, 1]$. By Lévy's characterization, M_t is a Brownian motion, $M_t = \sigma B_t$.

Thus

$$X_t = \gamma t + \sigma_t B_t + J_t$$

where $\gamma = \mu + \lambda\bar{\sigma}$. Now J_t is a finite variation martingale, so by Propn. 6.1 (assuming some integrability conditions) $\mathbb{E}[B_t J_t] = \mathbb{E}[\sum_{s \leq t} \Delta M_s \Delta J_s] = 0$, because B is continuous. This suggests that B, J may be independent, in which case the characteristic function of X_t is given by

$$\exp(i\mu t - \frac{1}{2}\sigma^2 t + \int e^{iux} - 1 \; \nu(dx))$$

where $\nu(dx) = \lambda\sigma(dx)$.

However, the initial assumption — that we can ignore jumps of amplitude $< \varepsilon$ — is not harmless. The sum of small jumps may be infinite! Nevertheless, some hard analysis shows that, in the limit $\varepsilon \to 0$, the characteristic function of X_t remains more or less in the above form. We do, however, need to compensate for the fact that the small jumps may add up to infinity by subtracting some drift, as follows: Arbitrarily assign jumps with amplitude ≤ 1 to be "small", and jumps with amplitude > 1 to be "large" — there is nothing special about the constant 1; any other $c > 0$ would do. Then, though the sum of the small jumps $\int_{|x| \leq 1} x \; N_t(dx)$ may be infinite, it turns out that the *compensated* integral $\int_{|x| \leq 1} x \; \tilde{N}_t(dx) = \int_{|x| \leq 1} x \; [N_t(dx) - t\nu(dx)]$ is guaranteed to be finite: $\int_{|x| \leq 1} x \; t\nu(dx)$ is, roughly, the expected sum of small jumps by time t, and subtracting the expected sum from the actual sum leaves us with something finite.

Without further ado, here is the statement of the Lévy–Ito Decomposition Theorem. Bretagnolle's[**?**] proof of this theorem may be found in Appendix G.

Theorem 7.1 (Lévy–Ito Decomposition)*Let X be a d–dimensional Lévy process. Then X has decomposition*

$$X_t = \gamma t + B_t + \int_{|x|\leq 1} x\ [N_t(dx) - t\nu(dx)] + \int_{|x|>1} x\ N_t(dx)$$

Here, $\gamma := \mathbb{E}[X_t - \int_{|x|>1} x\ N_t(dx)] \in \mathbb{R}^d$.

B_t is a centered Brownian motion with covariance matrix A.

The process $\int_{|x|\leq 1} x\ [N_t(dx) - t\nu(dx)]$ is a martingale independent of B_t.

For each $B \in \mathcal{B}(\mathbb{R}^d)$ bounded away from zero and each $f \cdot I_B \in \mathcal{L}^2(\nu)$, B_t is independent of $\int_B f(x)\ N_t(dx)$.

Furthermore, $\int_{B_1} f_1\ N_t(dx)$ and $\int_{B_2} f_2\ N_t(dx)$ are independent whenever B_1, B_2 are disjoint Borel sets which are bounded away from zero (assuming $f_1 \cdot I_{B_1}, f_2 \cdot I_{B_2} \in \mathcal{L}^2(\nu)$).

The measure ν is a Lévy measure, i.e. satisfies $\int_{\mathbb{R}^d} |x|^2 \wedge 1\ \nu(dx) < \infty$, and $\nu\{0\} = 0$.

□

The structure of the Lévy exponent (in the characteristic function) of a Lévy process is now apparent:

Corollary 7.2 *Let X be a d–dimensional Lévy process. Then*

$$\mathbb{E}[e^{i\langle u, X_t\rangle}] = \exp\left[t\left(i\langle u,\gamma\rangle - \frac{1}{2}\langle u, Au\rangle + \int_{\mathbb{R}^d} e^{i\langle u,x\rangle} - 1 - i\langle u,x\rangle I_{|x|\leq 1}\ \nu(dx)\right)\right]$$

Proof: By independence, the characteristic function of X_t is the product of the characteristic functions of $\gamma t + B_t$, $\int_{|x|>1} x\ N_t(dx)$ and $\int_{|x|\leq 1} x\ [N_t(dx) - t\nu(dx)]$.

Now $\mathbb{E}[e^{i\langle u, \gamma t + B_t\rangle}] = \exp[t(i\langle u,\gamma\rangle - \frac{1}{2}\langle u, Au\rangle)]$.

Further, by Propn. 6.8, $\int_{|x|>1} x\ N_t(dx)$ is a compound Poisson process, with characteristic function $\exp(t\int_{|x|>1} e^{i\langle u,x\rangle} - 1\ \nu(dx))$.

Finally, $\int_{|x|\leq 1} x\ [N_t(dx) - t\nu(dx)]$ is a limit of compensated compound Poisson processes $\int_{\frac{1}{n}<|x|\leq 1} x\ [N_t(dx) - t\nu(dx)]$, each of which has characteristic function

$$\exp\left(t\int_{\frac{1}{n}<|x|\leq 1} e^{i\langle u,x\rangle} - 1\ \nu(dx)\right)\exp\left(i\langle u, t\int_{\frac{1}{n}<|x|\leq 1} x\ \nu(dx)\right)$$

The limit of this sequence of characteristic functions is clearly

$$\exp\left(t\int_{|x|\leq 1} e^{i\langle u,x\rangle} - 1 - i\langle u,x\rangle\ \nu(dx)\right)$$

The result now follows easily by multiplication.

⊣

As an immediate consequence, we have the *Lévy–Khinchin Formula*. This result is often established prior to proving the Lévy–Ito Decomposition Theorem (cf., e.g., Sato[**?**]), but here it follows as a by–product:

Theorem 7.3 (Lévy–Khinchin)

(a) If μ is an infinitely divisible distribution on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, then its characteristic function has the form

$$\hat{\mu}(u) = \exp\left[i\langle u, \gamma\rangle - \frac{1}{2}\langle u, Au\rangle + \int_{\mathbb{R}^d} e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle I_{|x|\leq 1}\ \nu(dx)\right] \qquad (*)$$

where A is a symmetric non–negative definite $d \times d$–matrix, $\gamma \in \mathbb{R}^d$ and ν is a a Lévy measure on $\mathbb{R}^d$ (i.e. a measure satisfying $\nu\{0\} = 0$ and $\int_{\mathbb{R}^d} |x|^2 \wedge 1\ \nu(dx) < \infty$).

(b) Conversely, given a symmetric non–negative definite $d \times d$–matrix A, a Lévy measure ν on $\mathbb{R}^d$ and $\gamma \in \mathbb{R}^d$, then there is an infinitely divisible distribution μ on $\mathbb{R}^d$ whose characteristic function is given by $(*)$.

(c) The representation of $\hat{\mu}$ by A, ν and γ is unique.

Proof: (a) Let μ be an infinitely divisible distribution on $\mathbb{R}^d$. By Thm. 2.10, there exists a Lévy process X such that μ is the distribution of X_1. The result follows directly from the Lévy–Ito decomposition theorem and its corollary.

(b) For $n \in \mathbb{N}$, let

$$\varphi_n(u) = \exp\left[i\langle u, \gamma\rangle - \frac{1}{2}\langle u, Au\rangle + \int_{|x|>\frac{1}{n}} e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle I_{|x|\leq 1}\ \nu(dx)\right]$$

Note that because ν is a Lévy measure, we have $\nu(\{|x| > \frac{1}{n}\}) < \infty$. It follows that φ_n is the characteristic function of the convolution of a Gaussian distribution and a compound Poisson distribution, and thus infinitely divisible. The limit in distribution of these is therefore also infinitely divisible. This limit clearly has characteristic function $(*)$.

(c) Any two Lévy processes given by the same infinitely divisible distribution are identical in law. ν determines the law of the jumps of the associated Lévy process, and must be unique. Removing the jumps, A determines the quadratic variation of the associated continuous process, and must therefore be unique. It now follows easily that γ is unique also.

$\dashv$

Finally, for stochastic integration, it is important to note that Lévy processes are semimartingales:

Theorem 7.4 *Every Lévy process is a semimartingale.*

Proof: we can write $X_t = M_t + A_t$ where

$$M_t = B_t + \int_{|x|\leq 1} x\ [N_t(dx) - t\nu(dx)] \qquad \text{is a martingale}$$

and

$$A_t = \gamma t + \int_{|x|>1} x N_t(dx) \qquad \text{is a process of locally finite variation}$$

(because A_t has a.s. only finitely many jumps in each bounded interval.)

$\dashv$

8 Sample Path Properties of Lévy Processes

This section is taken mainly from Sato[**?**]. Let X be a Lévy process on $\mathbb{R}^d$ with generating triplet (A, γ, ν). Also, for $0 < a < b)$, let $D(a,b) \subseteq \mathbb{R}^d$ be the disc $D(a,b) = \{x \in \mathbb{R}^d : a < |x| \leq b\}$. Let J be the jump measure of X, i.e.

$$J(A,\omega) = \#\{t : (t, \Delta X_t) \in A\} \qquad \text{for } A \in \mathcal{B}((0,\infty) \times \mathbb{R}^d \backslash \{0\})$$

and let $N_t(B) = J((0,t] \times B)$. Also define, for $\varepsilon > 0$

$$\begin{aligned} J_\varepsilon(t,\omega) &= J((0,t] \times D(\varepsilon,\infty)) = N_t(\{|x| > \varepsilon\}) \\ X_\varepsilon(t,\omega) &= \int_{(0,t]\times D(\varepsilon,\infty)} x\ J(d(s,x),\omega) = \int_{D(\varepsilon,\infty)} x\ N_t(dx) \\ J(t,\omega) &= J((0,t] \times D(0,\infty)) = \lim_{\varepsilon \downarrow 0} J_\varepsilon(t,\omega) \end{aligned}$$

Proposition 8.1 *(Continuity) Sample paths of X are continuous iff $\nu = 0$.*

Proof: Note that J is a random measure with intensity measure $\lambda \times \nu$. $J_\varepsilon(t)$ (the number of jumps with amplitude $> \varepsilon$ by time t) has mean $\mathbb{E}[J_\varepsilon(t)] = \mathbb{E}[\int_{(0,t]\times D(\varepsilon,\infty)} J(d(s,x),\omega)] = t\nu(\{|x| > \varepsilon\})$. Hence the no. of jumps is zero a.s. iff $\nu = 0$ (using the fact that if $\xi \geq 0$ a.s., then $\xi = 0$ a.s. iff $\mathbb{E}[\xi] = 0$).

$\dashv$

Proposition 8.2 *(Jumping times)*

(a) If $\nu(\mathbb{R}^d) = \infty$, then a.s. jumping times are countable and dense in $[0,\infty)$.

(b) If $0 < \nu(\mathbb{R}^d) < \infty$, then a.s. jumping times are countably infinite in increasing order. Moreover, the interarrival period between jumps is exponentially distributed with mean $1/\nu(\mathbb{R}^d)$.

Proof: Any cádlág process, and hence any Lévy process, can have at most countably many jumps. Let τ_ε be the first time t that $|\Delta X_t| > \varepsilon$. Clearly $\tau_\varepsilon(\omega) \leq t$ iff $N_t(\{|x| > \varepsilon\}) \geq 1$. So

$$\mathbb{P}(\tau_\varepsilon(\omega) \leq t) = 1 - \mathbb{P}\big(N_t(\{|x| > \varepsilon\}) = 0\big) = 1 - e^{-t\nu(\{|x|>\varepsilon\})}$$

because N_t is a Poisson random measure with intensity measure $t\nu$. It follows that if $c_\varepsilon := \nu\{|x| > \varepsilon\} > 0$, then τ_ε is exponential with mean $1/c_\varepsilon$. (If $c_\varepsilon = 0$, then X has no jumps of size $> \varepsilon$.)

Now suppose first that $\nu(\mathbb{R}^d) = \infty$., then for any $t > 0$, we have

$$\lim_{\varepsilon \downarrow 0} \mathbb{P}(\tau_\varepsilon \leq t) = 1 - \lim_{\varepsilon \downarrow 0} e^{-t\nu(\{|x|>\varepsilon\})} = 1$$

It follows that $\lim_{\varepsilon \downarrow 0} \tau_\varepsilon = 0$ a.s. (as τ_ε is decreasing in ε, and bounded below by 0, hence convergent at $\varepsilon \downarrow 0$). It follows that, a.s. $t = 0$ is a cluster point of jump times of the path $X_t(\omega)$, i.e. there is a set $\Omega_0 \subseteq \Omega$ such that $\mathbb{P}(\Omega_0) = 1$, and such that for any $\omega \in \Omega_0$, the time $t = 0$ is a limit of jump times of $X_t(\omega)$. By the strong Markov property, any $u \geq 0$ is a cluster point of jump times also, i.e. for any $u \geq 0$ there is a set $\Omega_u \subseteq \Omega$ such that $\mathbb{P}(\Omega_u) = 1$,

and such that for any $\omega \in \Omega_u$, the time $t = 0$ is a right limit of jump times of $X_t(\omega)$. Let $\Omega' = \bigcap_{q\in\mathbb{Q}^+} \Omega_q$. Then for every $\omega \in \Omega'$, every rational time is a cluster point of jump times of $X_t(\omega)$, and $\mathbb{P}(\Omega') = 1$. Hence the set of jump times is dense

Next, assume instead that $0 < \nu(\mathbb{R}^d) < \infty$. Then $J(t)$ is a Poisson with mean $t\nu(\mathbb{R}^d) < \infty$, and $J(t) < \infty$ a.s. It follows that the jump times of X are enumerable in increasing order, and the strong Markov property guarantees that there are infinitely many jumps. Moreover, the first jump time is $\tau = \lim_{\varepsilon\downarrow 0} \tau_\varepsilon$ is exponential with mean $\lim_{\varepsilon\downarrow 0} \frac{1}{\nu(\{|x|>\varepsilon\})} = \frac{1}{\nu(\mathbb{R}^d)}$, as is easily seen by looking at the characteristic functions. By the strong Markov property, the same holds for all interarrival times.

$\dashv$

Next, we introduce the notion of *drift*: Recall that, by the Lévy–Ito decomposition Theorem, we may decompose a Lévy process as follows:

$$X_t = \gamma t + B_t + \int_{|x|\leq 1} x[N_t(dx) - t\nu(dx)] + \int_{|x|>1} xN_t(dx)$$

The term $\int_{|x|\leq 1} x[N_t(dx) - t\nu(dx)]$ includes a compensation term, to take care of the possibility that $\int_{|x|\leq 1} x\ N_t(dx)$ might be infinite. However, N_t is a Poisson random measure with intensity measure $t\nu$, and so $\mathbb{E}[\int_{|x|\leq 1} x\ N_t(dx)] = t\int_{|x|\leq 1} x\ \nu(dx)$. It follows that if $\int_{|x|\leq 1} |x|\ \nu(dx) < \infty$, then also $\int_{|x|\leq 1} x\ N_t(dx) < \infty$ a.s., and hence the compensation term is not necessary. Thus, in the case that $\int_{|x|\leq 1} |x|\ \nu(dx) < \infty$, we may write the Lévy–Ito decomposition of X as

$$X_t = \gamma_0 t + B_t + \int_{\mathbb{R}^d} x\ N_t(dx) \qquad \text{where } \gamma_0 = \gamma - \int_{|x|\leq 1} x\ \nu(dx)$$

The vector γ_0 is called the *drift* of the Lévy process X. Note that a Lévy process has drift only if $\int_{|x|\leq 1} |x|\ \nu(dx) < \infty$. It is now easy to see that the characteristic function of such a Lévy process

$$\mathbb{E}[e^{i\langle u, X_t\rangle}] = \exp\left[t\Big(i\langle u, \gamma_0\rangle - \frac{1}{2}\langle u, Au\rangle + \int_{\mathbb{R}^d} e^{i\langle u,x\rangle} - 1\ \nu(dx)\Big)\right]$$

Sample paths of a Lévy process with drift are easy to visualize: Each path is a sum of a linear drift, a Brownian motion and jumps. Note that if $\nu(\mathbb{R}^d) < \infty$, then X has a drift. However, there are also Lévy processes with drift for which $\nu(\mathbb{R}^d) = \infty$ (e.g. the gamma process).

Following Sato[**?**], we define a Lévy process X on $\mathbb{R}^d$ with generating triplet (A, ν, γ) to be

Of type A: if $A = 0$ and $\nu(\mathbb{R}^d) < \infty$

Of type B: if $A = 0, \nu(\mathbb{R}^d) = \infty$ and $\int_{|x|\leq 1} |x|\ \nu(dx) < \infty$

Of type C: else, i.e. if $A \neq 0$ or $\int_{|x|\leq 1} |x|\ \nu(dx) = \infty$.

These types exhaust the possibilities. Note that types A and B are Lévy processes with drift. Moreover, type A consists of processes that have jump times increasing in order, whereas types B,C have jump times dense.

Proposition 8.3 *(Increasingness) A one–dimensional Lévy process X is increasing iff it $A = 0$, $\nu(-\infty, 0) = 0$, $\int_{(0,1]} x\ \nu(dx) < \infty$ and has drift $\gamma_0 \geq 0$.*

A *subordinator* is an increasing (and thus 1–dimensional) Lévy process.

Proof: $\Leftarrow$: First note that Since $\mathbb{E}[J((0,t] \times (-\infty, 0))] = t\nu(-\infty, 0) = 0$, we see that X has a.s. no negative jumps, so that

$$X_t = \gamma_0 t + \int_{(0,\infty)} x\ N_t(dx) \qquad \text{a.s.}$$

It follows immediately that X is increasing.

$\Rightarrow$: If X is increasing, it cannot have negative jumps, form which it follows that $\nu(-\infty, 0) = 0$. Next, note that each process $X_t - X_\varepsilon(t)$ is increasing also, being X with some of its jumps removed. In particular, $X_t - X_\varepsilon(t) \geq 0$. Define

$$\tilde{X}_t = \lim_{\varepsilon \downarrow 0} X_\varepsilon(t) = \int_{(0,\infty)} x\ N_t(dx)$$

This limit exists, because $X_\varepsilon(t)$ is increasing as $\varepsilon \downarrow 0$, and bounded by X_t, as X is increasing. Now each $X_\varepsilon(t)$ is a compound Poisson process with Laplace transform

$$\begin{aligned}
\mathbb{E}[e^{-uX_\varepsilon(t)}] &= \exp\left[t \int_{(\varepsilon,\infty)} e^{-ux} - 1\ \nu(dx)\right] \\
&= \exp\left[t \int_{(\varepsilon,\infty)} e^{-ux} - 1 + uxI_{(0,1]}(x)\ \nu(dx) - tu \int_{(\varepsilon,1]} x\ \nu(dx)\right]
\end{aligned}$$

where $u > 0$. By the dominated convergence theorem, $\mathbb{E}[e^{-uX_\varepsilon(t)}] \to \mathbb{E}[e^{-u\tilde{X}_t}]$ as $\varepsilon \downarrow 0$. Since $\int_{(\varepsilon,\infty)} e^{-ux} - 1 + uxI_{(0,1]}(x)\ \nu(dx)$ tends to $\int_{(0,\infty)} e^{-ux} - 1 + uxI_{(0,1]}(x)\ \nu(dx) < \infty$ and $\int_{(\varepsilon,1]} x\ \nu(dx)$ tends to $\int_{(0,1]} x\ \nu(dx)$ as $\varepsilon \downarrow 0$. Since $\mathbb{E}[e^{-u\tilde{X}_t}] > 0$ and $\int_{(0,\infty)} e^{-ux} - 1 + uxI_{(0,1]}(x)\ \nu(dx) < \infty$, we must have $\int_{(0,1]} x\ \nu(dx)$ also, i.e. X has , which we denote γ_0. It follows that we may write $X_t = \gamma_0 t + B_t + \int_{(0,\infty)} x\ N_t(dx) = \gamma_0 t + B_t + \tilde{X}_t$. Since $X_t - \tilde{X}_t \geq 0$, it follows that $B_t = 0$, i.e. that $A = 0$.

$\dashv$

Recall that if f is a function $f : [0, \infty) \to \mathbb{R}^d$, and if $0 \leq t_1 < t_2 < \infty$, then the variation $V(f; (t_1, t_2])$ of f over the interval $(t_1, t_2]$ is defined by

$$V(f; (t_1, t_2]) = \sup_\Delta \sum_{j=1}^n |f(s_j) - f(s_{j-1})|$$

Proposition 8.4 *If X is of type A or B, then sample paths of X are a.s. of finite variation on $(0, t]$. If X is of type C, then sample paths of X are a.s. of unbounded variation on $(0, t]$.*

Proof: First suppose that X is of type A or B. Then $\int_{D(0,1]} |x|\ \nu(dx) < \infty$. Define

$$U_t = \int_{D(0,\infty)} |x|\ N_t(dx) = \sum_{0<s\leq t} |X(s) - X(s-)|$$

We first claim that $U_t < \infty$ a.s. Indeed, we have $U_t = \int_{D(0,1]} |x| \; N_t(dx) + \int_{D(1,\infty)} |x| \; N_t(dx)$. The second term is obviously finite. The first term has expectation $\mathbb{E}[\int_{D(0,1]} |x| \; N_t(dx)] = t\int_{D(0,1]} |x| \; \nu(dx)$, which is finite also. Hence $U_t < \infty$. Since X_t is of type A or B, we have

$$X_t = \gamma_0 t + \int_{D(0,\infty)} x N_t(dx) = \gamma_0 t + \sum_{0<s\leq t} (X(s) - X(s-))$$

Clearly, $V(X;(0,t]) = |\gamma_0|t + U_t < \infty$, i.e. X is of finite variation on $(0,t]$.

Next, suppose that X is of type C. Then either $A \neq 0$ or $\int_{D(0,1]} |x| \; \nu(dx) = \infty$. Assume first that $\int_{D(0,1]} |x| \; \nu(dx) = \infty$. Define U_t as above, and note that $U_t = \lim_{\varepsilon\downarrow 0} U_\varepsilon(t)$, where $U_\varepsilon(t) = \int_{D(\varepsilon,\infty)} |x| \; N_t(dx)$. Each $U_\varepsilon(t)$ is a compound Poisson process with Laplace transform

$$\begin{aligned}
\mathbb{E}[e^{-uU_\varepsilon(t)}] &= \exp\left[t\int_{D(\varepsilon,\infty)} e^{-u|x|} - 1 \; \nu(dx)\right] \\
&= \exp\left[t\int_{D(\varepsilon,\infty)} e^{-u|x|} - 1 + u|x|I_{D(0,1]}(x) \; \nu(dx) - tu\int_{D(\varepsilon,1]} |x| \; \nu(dx)\right]
\end{aligned}$$

for $u > 0$. If we assume that $\int_{D(0,1]} |x| \; \nu(dx) = \infty$, then $\mathbb{E}[e^{-uU_\varepsilon(t)}] \to 0$ as $\varepsilon \downarrow 0$. Thus $U_t = \infty$ a.s. for any $t > 0$. Since clearly $V(X;(0,t]) \geq U_t$, we see that X is a.s. of unbounded variation on $(0,t]$, for any $t > 0$.

It remains to show that if X of unbounded variation on $(0,t]$ if it is of type C with $A \neq 0$ and $\int_{D(0,1]} |x| \; \nu(dx) < \infty$. In that case, X has drift: $X_t = \gamma_0 t + B_t + \int_{D(0,\infty)} x \; N_t(dx)$. The process $\gamma_0 t + \int_{D(0,\infty)} x \; N_t(dx)$ is of bounded variation, whereas B_t is (as is well–known) of unbounded variation on $(0,t]$. The result follows.

$\dashv$

Proposition 8.5 *Suppose that X is of type A or B. Let $V_t = V(X;(0,t])$ be the variation of X. Then V_t is a subordinator with Laplace transform*

$$\mathbb{E}[e^{-uV_t}] = \exp\left[t\int_{\mathbb{R}^d} e^{-u|x|} - 1 \; \nu(dx) - |\gamma_0|ut\right]$$

for $u > 0$. Hence V_t has drift $|\gamma_0|$ and Lévy measure defined by $\nu_0(B) = \int_{\mathbb{R}^d} I_B(|x|) \; \nu(dx)$ for $B \in \mathcal{B}(\mathbb{R})$.

Proof: As in the proof of Propn. 8.4, we have $V_t = |\gamma_0|t + U_t$, where

$$U_t = \int_{D(0,\infty)} |x| \; N_t(dx) = \sum_{0<s\leq t} |X(s) - X(s-)|$$

It is not hard to see that V_t is cádlág with independent increments. Define $U_\varepsilon(t) = \int_{D(\varepsilon,\infty)} |x| \; N_t(dx)$, so that $U_\varepsilon(t)$ is compound Poisson with

$$\mathbb{E}[e^{-u(U_\varepsilon(t) - U_\varepsilon(s))}] = \exp\left[(t-s)\int_{D(\varepsilon,\infty)} e^{-u|x|} - 1 \; \nu(dx)\right]$$

In the limit $\varepsilon \downarrow 0$, we obtain

$$\mathbb{E}[e^{-u(V_t-V_s)}] = \exp[u|\gamma_0|(t-s)] \exp\left[(t-s)\int_{D(0,\infty)} e^{-u|x|} - 1 \; \nu(dx)\right]$$

and it follows that V_t is a Lévy process with the stated Laplace transform. Since $\int_{\mathbb{R}^d}(e^{-u|x|} - 1)\, \nu(dx) = \int_{\mathbb{R}}(e^{-uy} - 1)\, \nu_0(dy)$, ν_0 is the Lévy measure of V_t.

$\dashv$

A Technical Results about Infinitely Divisible Distributions

The following are taken from Sato[**?**]:

Lemma A.1 *If μ is infinitely divisible, then $\hat{\mu}(u) \neq 0$ for any $u \in \mathbb{R}^d$.*

Proof: First note that if ν is a probability distribution, then $|\hat{\nu}|^2$ is a characteristic function: Indeed, if we define $\tilde{\nu}(B) = \nu(-B)$, then $|\hat{\nu}|^2 = \widehat{\nu * \tilde{\nu}}$.

Suppose that μ is infinitely divisible, and that $\mu = \mu_n^n$ for $n \in \mathbb{N}$. Then $|\hat{\mu}_n(u)|^2 = |\hat{\mu}(u)|^{2/n}$. Note that $|\hat{\mu}(u)|^{2/n} \to 1$ if $\hat{\mu}(u) \neq 0$, and that $|\hat{\mu}(u)|^{2/n} \to 0$ if $\hat{\mu}(u) = 0$. Let

$$\varphi(u) = \lim_n |\hat{\mu}_n(u)|^2 = \begin{cases} 1 & \text{if } \hat{\mu}(u) \neq 0 \\ 0 & \text{if } \hat{\mu}(u) = 0 \end{cases}$$

Since $\hat{\mu}$ is continuous, and since $\hat{\mu}(0) = 1$, we must have $\hat{\mu} \neq 0$ on some neighbourhood of 0. It follows that φ is continuous on a neighbourhood of 0. Since φ is also a pointwise limit of characteristic functions, it is itself a characteristic function, and thus continuous, by Thm. 2.3. Hence $\varphi = 1$ identically, and thus $\hat{\mu}(u)$ cannot equal 0 at any point.

$\dashv$

Lemma A.2 *Suppose that $\varphi : \mathbb{R}^d \to \mathbb{C}$ is a continuous function with $\varphi(0) = 1$ and $\varphi(u) \neq 0$ for any $u \in \mathbb{R}^d$.*

(a) There is a unique continuous $f : \mathbb{R}^d \to \mathbb{C}$ such that $\varphi(u) = e^{f(u)}$.
We write $f(u) = \log \varphi(u)$ and call it the distinguished logarithm *of φ.*

(b) For any $n \in \mathbb{N}$, there is a unique continuous $g_n : \mathbb{R}^d \to \mathbb{C}$ such that $g_n(0) = 1$ and $g_n(u)^n = \varphi(u)$. In fact, $g_n(u) = e^{f(u)/n}$.
We write $g_n(u) = \varphi(u)^{\frac{1}{n}}$, and call it the distinguished n^{th} root *of φ.*

Remarks A.3 Note that $\varphi(u_1) = \varphi(u_2)$ need not imply that $f(u_1) = f(u_2)$. For example, if $\varphi : \mathbb{R} \to \mathbb{C} : u \mapsto e^{2\pi i u}$, then $f(u) = 2\pi i u$. Now $\varphi(0) = \varphi(1)$, but $f(0) \neq f(1)$. Hence f is not simply the composition of a branch of the complex logarithm with φ.

□

Proof: This proof requires some knowledge of complex analysis.
(a) Let $\log^{(\mathbb{C})} z$ denote the complex logarithm, a multifunction:

$$\log^{(\mathbb{C})} z = \log |z| + i \arg z$$

where $\arg z$ is any argument of z, determined up to a multiple of 2π. For a fixed $u \in \mathbb{R}^d$, the map $\varphi(tu), 0 \leq t \leq 1$ traces a curve in $\mathbb{C} \backslash \{0\}$ starting at 1. Let $h_u(t), 0 \leq t \leq 1$ be the unique branch of $\log^{(\mathbb{C})} \varphi(tu)$ which has $h_u(0) = 0$ and such that h_u is continuous in t. Now define f by

$$f(u) = h_u(1)$$

Clearly $f(0) = h_0(1) = 0$. Also, $e^{f(u)} = e^{h_u(1)} = \varphi(u)$. It therefore remains to show that (i) f is *continuous*, and (ii) f is *unique*.

First, we show that f is continuous. Fix $u_0 \in \mathbb{R}^d$, and let $u\mathbb{R}^d$ with $u \neq u_0$. We want to show that $f(u)$ is close to $f(u_0)$ when u is close to u_0. Define $\gamma_{u_0,u} : [0,3] \to \mathbb{R}^d$ by

$$\gamma_{u_0,u}(t) = \begin{cases} tu_0 & \text{if } 0 \leq t \leq 1 \\ (t-1)u + (2-t)u_0 & \text{if } 1 \leq t \leq 2 \\ (3-t)u & \text{if } 2 \leq t \leq 3 \end{cases}$$

Thus $\gamma_{u_0,u}$ traces a triangle from 0 to u_0 to u to 0 — a closed curve in $\mathbb{R}^d$. Hence $\varphi(\gamma(t))$ traces a closed curve in $\mathbb{C}\backslash\{0\}$ starting and ending at 1. Let $\theta_{u_0,u}(t), 0 \leq t \leq 3$ be the branch of $\arg \varphi(\gamma_{u_0,u}(t))$ which is continuous in t and has $\theta_{u_0,u}(0) = 0$. Now the set $\{\varphi(tu_0) : t \in [0,1]\}$ is a compact subset of $\mathbb{C}$ which does not contain 0, and hence is a strictly positive distance away from zero. Furthermore, $\max_{t\in[0,1]} |\varphi(tu) - \varphi(tu_0)| \to 0$ as $u \to u_0$, because continuous functions on compacts are uniformly continuous. Now the closed curve $\varphi(\gamma_{u_0,u}(t)), 0 \leq t \leq 3$ is obtained by gluing three curves together, namely the curve $\varphi(tu_0), 0 \leq t \leq 1$, the curve $\varphi(tu + (1-t)u_0), 0 \leq t \leq 1$, and the curve $\varphi(tu), 0 \leq t \leq 1$. Thus if we take u sufficiently close to u_0, then the third curve is very close to the first curve, but traces it *in reverse.* If u is sufficiently close to u_0, the region enclosed by the closed curve must exclude the origin (as the first part of the curve is a positive distance away from the origin, and the other parts lie very close to it). Thus if u is sufficiently close to u_0, the winding number of the closed curve $\varphi(\gamma_{u_0,u}(t)), 0 \leq t \leq 3$ is zero. It follows that $\theta_{u_0,u}(3) = 0$ for all u in a neighbourhood of u_0. Now consider $f(u) = h_u(1)$, where h_u is the unique continuous branch of $\log^{(\mathbb{C})} \varphi(tu) = \log|\varphi(tu)| + i \arg \varphi(tu)$ with $h_u(0) = 0$. Thus $\mathrm{Im} f(u)$ is an argument of $\varphi(u)$, and indeed $\mathrm{Im} f(u) = \theta_{u_0,u}(2)$ (for u in a neighbourhood of u_0). The continuity of $\theta_{u_0,u}$ implies that $\mathrm{Im} f(u)$ is close to $\mathrm{Im} f(u_0)$ whenever u is close to u_0, and thus that f is continuous (as $\mathrm{Re} f(u) = \log|\varphi(u)|$ is obviously continuous).

Next we show that f is unique: Suppose that $\tilde{f}$ is another continuous map from $\mathbb{R}^d$ to $\mathbb{C}$ which satisfies $e^{\tilde{f}(u)} = \varphi(u)$ and $\tilde{f}(0) = 0$. Then $t \mapsto \tilde{f}(tu), 0 \leq t \leq 1$ is clearly a continuous branch of $\log^{(\mathbb{C})} \varphi(tu)$, with $e^{\tilde{f}(0)} = 1$. It follows that $\tilde{f}(tu) = h_u(t)$ (by uniqueness of h_u), and thus that $\tilde{f} = f$.

(b) Clearly $g_n(u) = e^{f_n(u)/n}$ satisfies all the requirements. Uniqueness, follows as above, noting that the complex n^{th} root of z is $|z|^{\frac{1}{n}} e^{\frac{i}{n} \arg z}$, a multifunction which must have a unique continuous branch h_n satisfying $h_n(1) = 1$.

$\dashv$

Corollary A.4 *If μ is an infinitely divisible distribution on $\mathbb{R}^d$, then there is, for each $n \in \mathbb{N}$, a unique μ_n such that $\mu_n^n = \mu$. Moreover, $\hat{\mu}_n = \hat{\mu}^{\frac{1}{n}}$.*

Proof: By Lemma A.1, $\hat{\mu}(u) \neq 0$ for any $u \in \mathbb{R}^d$. Hence, by Lemma A.2, there is a unique continuous g_n with $g_n(0) = 1$ such that $\hat{\mu}(u) = g_n(u)^n$ (i.e. $g_n = \hat{\mu}^{\frac{1}{n}}$). Now since μ is infinitely divisible, there exists a probability measure μ_n with $\mu = \mu_n^n$, i.e. $\hat{\mu} = \hat{\mu}_n^n$. Uniqueness implies that $\hat{\mu}_n = g_n$. If ν_n is another measure such that $\nu_n^n = \mu$, then also $\hat{\nu}_n = g_n$, i.e. $\hat{\mu}_n = \hat{\nu}_n$, and so $\mu_n = \nu_n$, by Theorem 2.3 (Uniqueness).

$\dashv$

Note that it isn't generally the case that $\mu^n = \nu^n$ implies $\mu = \nu$. This does follow if μ or ν is infinitely divisible, as is easy to see.

Proposition A.5 *If μ_n are infinitely divisible (for $n \in \mathbb{N}$) and if $\mu_n \to \mu$ weakly, then μ is infinitely divisible.*

Proof: We first show that $\hat{\mu}(u) \neq 0$ for any $u \in \mathbb{R}^d$. Now $\hat{\mu}_n(u) \to \hat{\mu}(u)$, by Glivenko's Theorem, and so $|\hat{\mu}_n(u)|^{2/k} \to |\hat{\mu}(u)|^{2/k}$ for $k \in \mathbb{N}$ as $n \to \infty$. Each $|\hat{\mu}_n|^{2/k}$ is itself a characteristic function, namely the k^{th} root of the transform of the convolution $\mu_n * \tilde{\mu}_n$ of two infinitely divisible distributions (where $\tilde{\mu}_n(B) = \mu_n(-B)$). Since $|\hat{\mu}(u)|^{2/k}$ is continuous, it is a characteristic function, by Thm. 2.3 (Lévy continuity), and so $|\hat{\mu}(u)|^2 = (|\hat{\mu}(u)|^{2/k})^k$ is the characteristic function of an infinitely divisible distribution. In particular, by Lemma A.1, $|\hat{\mu}(u)|^2 \neq 0$ for any $u \in \mathbb{R}^d$, and thus $\hat{\mu}(u) \neq 0$ for any u. By Thm. 2.3 (Glivenko), $\hat{\mu}_n \to \hat{\mu}$ uniformly on compacts, and thus, by construction of the distinguished logarithm, we see that $\log \hat{\mu}_n \to \log \hat{\mu}$ uniformly on compacts. In particular, $\hat{\mu}_n(u)^{1/k} \to \hat{\mu}^{1/k}(u)$ (recall $\hat{\mu}^{1/k}(u) = e^{\frac{1}{k}\log \hat{\mu}(u)}$). Since $\hat{\mu}^{1/k}$ is continuous, it is a characteristic function, by Thm. 2.3 (Lévy continuity), and hence μ is infinitely divisible.

$\dashv$

B Proof of Existence of Lévy Processes

We follow Sato[**?**] for the proof of the next theorem:

Theorem B.1 *Let μ be an infinitely divisible distribution on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Then there is a Lévy process in law X such that μ is the distribution of X_1.*
Moreover, if X' is another such Lévy process, then X, X' are identical in law.

Proof: We apply Kolmogorov's existence theorem: Let $\Omega = (\mathbb{R}^d)^{[0,\infty)}$ be the set all functions $\omega = \omega(t)$ from $[0, \infty)$ to $\mathbb{R}^d$, and let $X_t(\omega) = \omega(t)$ be the coordinate process. Equip Ω with the σ–algebra $\mathcal{F}$ generated by X, i.e. $\mathcal{F}$ is generated by the so–called *cylinder sets*, which are sets of the form

$$C = \{\omega : X_{t_0}(\omega) \in B_0, \dots, X_{t_n} \in B_n\}$$
$$\text{where } 0 \leq t_0 < t_1 < \cdots < t_n \text{ and } B_0, \dots, B_n \in \mathcal{B}(\mathbb{R}^d)$$

For such t_k, B_k, define

$$\begin{aligned} &\mu_{t_0,\dots,t_n}(B_0 \times \cdots \times B_n) \\ &= \int \dots \int \mu^{t_0}(dy_0)\, I_{B_0}(y_0) \mu^{t_1-t_0}(dy_1)\, I_{B_1}(y_0+y_1) \dots \mu^{t_n-t_{n-1}}(dy_n)\, I_{B_n}(y_0+\cdots+y_n) \end{aligned}$$

This defines $\mu_{t_0,\dots,t_n}$ on rectangles in $\mathcal{B}((\mathbb{R}^d)^{n+1})$. By standard measure–theoretic argument, $\mu_{t_0,\dots,t_n}$ can be extended to a unique measure on $\mathcal{B}((\mathbb{R}^d)^{n+1})$. It is easy to verify that the $\mu_{t_o,\dots,t_n}$ satisfy the so–called consistency condition of the Kolmogorov existence theorem, i.e. that when $B_k = \mathbb{R}^d$ we have

$$\mu_{t_0,\dots,t_n}(B_0 \times \cdots \times B_n) = \mu_{t_0,\dots,t_{k-1},t_{k+1}\dots,t_n}(B_0 \times \cdots \times B_{k-1} \times B_{k+1} \times B_n)$$

The Kolmogorov's existence theorem now guarantees the existence of a unique measure $\mathbb{P}$ on $\mathcal{F}$ which has the $\mu_{t_0,\dots,t_n}$ as its finite–dimensional distributions, i.e. for which

$$\mathbb{P}(X_{t_0} \in B_0, \dots X_{t_n} \in B_n) = \mu_{t_0,\dots,t_n}(B_0 \times \cdots \times B_n)$$

Thus

$$\mathbb{P}(X_t \in B) = \mu_t(B) = \int \mu^t(dy)\ I_B(y) = \mu^t(B)$$

i.e. X_t has distribution μ^t, and, in particular, X_1 has distribution μ.

We will now show that X_t is a Lévy process in law: It is straightforward to show that

$$\mathbb{E}[f(X_{t_0}, \dots X_{t_n})] = \int \dots \int f(y_0, y_0 + y_1, \dots, y_0 + \dots y_n) \mu^{t_0}(dy_0) \mu^{t_1 - t_0}(dy_1) \dots \mu^{t_n - t_{n-1}}(dy_n)$$

for every bounded measurable f. In particular, for $u_1, \dots, u_n \in \mathbb{R}^d$ we have

$$\mathbb{E}\left[e^{i \sum_{k=1}^n \langle u_j, X_{t_j} - X_{t_{j-1}} \rangle}\right] = \prod_{j=1}^n \int e^{i\langle u_j, y_j\rangle}\ \mu^{t_j - t_{j-1}}(dy_j)$$

Thus, with $n = 1$, we see that

$$\mathbb{E}\left[e^{i\langle u, X_t - X_s\rangle}\right] = \int e^{i\langle u, y\rangle} \mu^{t-s}(dy)$$

which proves that $X_t - X_s$ has distribution μ^{t-s}, and thus that X has stationary increments. Furthermore, we can now deduce that

$$\mathbb{E}\left[e^{i \sum_{k=1}^n \langle u_j, X_{t_j} - X_{t_{j-1}} \rangle}\right] = \prod_{j=1}^n \mathbb{E}\left[e^{i\langle u_j, X_{t_j} - X_{t_{j-1}}\rangle}\right]$$

which shows that X_t has independent increments.

It remains only to prove that X is continuous in probability. Now we have $\mu^t \to 0$ in distribution as $t \downarrow 0$, i.e. $X_t \to 0$ in distribution. Convergence in distribution to a constant implies convergence in probability, and thus $\lim_{t\downarrow 0} \mathbb{P}(|X_t| > \varepsilon) = 0$ for any $\varepsilon > 0$. But then

$$\lim_{s \to t} \mathbb{P}(|X_t - X_s| > \varepsilon) = \lim_{|s-t|\downarrow 0} \mathbb{P}(|X_{|s-t|}| > \varepsilon) = 0$$

by stationarity of increments.

$\dashv$

Theorem B.2 *If X is a Lévy process in law, then there is a unique*[2] *modification Y of X such that Y is càdlàg, and therefore a Lévy process.*

Proof:

For $u \in \mathbb{R}^d$, let $\varphi_t(u) = \mathbb{E}[e^{i\langle u, X_t\rangle}]$ be the characteristic function of X_t. (Then $\varphi_t(u) = e^{-t\psi(u)}$ for some continuous ψ, and hence φ_t is continuous.) For each $u \in \mathbb{R}^d$, the process

$$M_t^u := \frac{e^{i\langle u, X_t\rangle}}{\varphi_t(u)}$$

is a complex–valued $\mathcal{F}_t$–martingale (cf. Propn 3.1 Now it is well–know that every martingale has a càdlàg modification (cf. , e.g. Revuz and Yor[**?**] Thm II.2.9). If N_t^u is a càdlàg modification of M_t^u, then

$$\{\omega \in \Omega : \exists q \in \mathbb{Q}_+ (M_q^u \neq N_q^u\}$$

[2]unique up to indistinguishability.

is a null set, from which it follows that M_t^u is càdlàg on $\mathbb{Q}_+$. Thus for almost every ω, the map $q \mapsto M_q^u(\omega)$ is right–continuous with left limits (RCLL) on $\mathbb{Q}_+$. It follows that, for almost all ω, the maps $q \mapsto e^{i\langle u, X_q\rangle}$ is RCLL also on $\mathbb{Q}_+$. Hence for almost all ω, the map $q \mapsto e^{i\langle u, X_q\rangle}$ is RCLL on $\mathbb{Q}_+$ simultaneously for all $u \in \mathbb{Q}^d$.

We now show that the map $q \mapsto X_q(\omega)$ is RCLL on $\mathbb{Q}_+$. For suppose that this map doesn't have left limits on $\mathbb{Q}_+$. Then we can find $q \in \mathbb{Q}_+$ and $q_n \uparrow q$, $q'_n \uparrow q$ such that $\lim_{n\to\infty} X_{q_n}(\omega) = a \neq a' = \lim_{n\to\infty} X_{q'_n}$. We may then choose $u \in \mathbb{Q}^d$ such that $\langle u, a - a'\rangle \notin 2\pi i\mathbb{Z}$, and hence $\lim_n e^{i\langle u, X_{q_n}\rangle} \neq \lim_n e^{i\langle u, X_{q'_n}\rangle}$, which contradicts the fact that $q \mapsto e^{i\langle u, X_q\rangle}$ has left limits on $\mathbb{Q}_+$. A similar argument, with $q_n, q'_n \downarrow q$, shows that $q \mapsto X_q(\omega)$ is right–continuous for almost all ω.

Let N be the null set of ω's for which the map $q \mapsto X_q(\omega)$ is not càdlàg. Define, for $t \in \mathbb{R}_+$

$$Y_t(\omega) = \begin{cases} \lim_{q\downarrow t, q\in\mathbb{Q}} X_q(\omega) & \text{if } \omega \in N^c \\ 0 & \text{else} \end{cases}$$

Then Y is clearly càdlàg. Moreover, because X is continuous in probability Because $\mathcal{F}_t$ contains all the $\mathbb{P}$–null sets of $\mathcal{F}$ and is right–continuous, we see that Y_t is adapted to $\mathcal{F}_t$. Because X is continuous in probability), we have $X_q \to X_s$ in probability, and thus in distribution, as $q \downarrow t$. By the dominated convergence theorem, we therefore have

$$\mathbb{E}[e^{i\langle u, Y_t - X_t\rangle}] = \lim_{q\downarrow t, q\in\mathbb{Q}} \mathbb{E}[e^{i\langle u, X_q - X_t\rangle}] = 1$$

It follows that $\mathbb{P}(Y_t \neq X_t) = 0$, i.e. that Y is a modification of X. It is now immediate that Y has independent identically distributed increments also, and thus that Y is a Lévy process.

$\dashv$

Remarks B.3 The above proof uses martingale theory to prove the existence of a càdlàg version of a Lévy process in law. Sato provides a direct proof of a stronger result: Every additive process in law has a càdlàg version. The martingale proof cannot extend to this, because not every additive process is a semimartingale (e.g. every deterministic process is additive, but it is a semimartingale only if it is of bounded variation on compacts).

$\square$

Corollary B.4 *Let μ be an infinitely divisible distribution on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Then there is a Lévy process X such that μ is the distribution of X_1. Any two Lévy processes with this property are identical in law.*

C Right–Continuity of Filtration

Wherever filtrations are mentioned in a continuous–time context, it is almost always assumed that the *usual hypotheses* hold. Recall that a filtration $(\mathcal{F}_t)_{t\geq 0}$ is on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ said to satisfy the usual hypotheses if (i) $\mathcal{F}_0$ contains all the $\mathbb{P}$–null sets in $\mathcal{F}$, and (ii) $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$ (i.e. the filtration is *right–continuous*). A filtration $\mathcal{F}_t^o$ can always be augmented to satisfy the usual conditions:

(1.) First define $\bar{\mathcal{F}}_t^o = \sigma(\mathcal{F}_t^o \cup \mathcal{N})$, where $\mathcal{N}$ is the set of $\mathbb{P}$–null sets in $\mathcal{F}$. Note that $\bar{\mathcal{F}}_t^o = \{X \in \mathcal{F} : \exists A, B \in \mathcal{F}_t^o (A \subseteq X \subseteq B \quad \wedge \quad B \backslash A \in \mathcal{N})\}$.

(2.) Then define $\mathcal{F}_t = \bigcap_{s>t} \bar{\mathcal{F}}_t^o$.

Then $(\mathcal{F}_t)_t$ is the smallest filtration which contains $(\mathcal{F}_t^o)_t$ and satisfies the usual conditions. Thus to construct the augmented filtration $\mathcal{F}_t$ of $\mathcal{F}_t^o$, we must proceed in two steps. The main result of this section shows that if $\mathcal{F}_t^o$ is the natural filtration of a Lévy process X, i.e. if $\mathcal{F}_t^o = \sigma(X_s : s \leq t)$, then the second step is unnecessary: The filtration $\bar{\mathcal{F}}_t^o$ is already right–continuous (and so $\mathcal{F}_t = \bar{\mathcal{F}}_t^o$).

First note that if (M_t) is a $\mathcal{F}_t^o$–martingale, it is also a $\bar{\mathcal{F}}_t^o$–martingale: For if $\mathbb{E}[M_t|\mathcal{F}_s^o] = M_s$, then M_s is certainly also $\bar{\mathcal{F}}_s^o$–measurable. Moreover, if $X \in \bar{\mathcal{F}}_s^o$, and if $A, B \in \mathcal{F}_s^o$ such that $A \subseteq X \subseteq B$ and $B \backslash A \in \mathcal{N}$, then $\mathbb{E}[M_t; X] = \mathbb{E}[M_t; A] = \mathbb{E}[M_s; A] = \mathbb{E}[M_s; X]$. This shows that $\mathbb{E}[M_t|\bar{\mathcal{F}}_s^o] = M_s$ a.s.

As a consequence, if X_t is a Lévy process with natural filtration $\mathcal{F}_t^o$, then the martingales $M_t^u = \frac{e^{iuX_t}}{\varphi_t(u)}$ defined in Propn. 3.1 are also $\bar{\mathcal{F}}_t^o$–martingales. We need this in the proof of the next result.

Theorem C.1 *Let X be a Lévy process on $(\Omega, \mathcal{F}, \mathbb{P})$ with canonical filtration $(\mathcal{F}_t^o)_t$, and let $\bar{\mathcal{G}}_t = \sigma(\mathcal{F}_t^o \cup \mathcal{N})$ (where $\mathcal{N}$ is the collection of $\mathbb{P}$–null sets in $\mathcal{F}$). Then $(\mathcal{G}_t)_t$ is right–continuous.*

Proof: We follow Protter[**?**]: Let $\mathcal{G}_{t+} = \bigcap_{s>t} \mathcal{G}_s$. We must show that $\mathcal{G}_{t+} = \mathcal{G}_t$. We first show that

$$\mathbb{E}[e^{i\sum_{j=1}^n \langle u_j, X_{s_j}\rangle}|\mathcal{G}_{t+}] = \mathbb{E}[e^{i\sum_{j=1}^n \langle u_j, X_{s_j}\rangle}|\mathcal{G}_t] \qquad (*)$$

for all $0 \leq s_1, \ldots, s_n$ and all $u_1, \ldots u_n \in \mathbb{R}^d$. Now if $s_1, \ldots, s_n \leq t$, then equality is obvious, as both sides are equal to $e^{i\sum_{j=1}^n \langle u_j, X_{s_j}\rangle}$. It therefore remains to show that the result holds for $s_1, \ldots s_n > t$, and for notational simplicity, we shall prove it for the case $n = 2$, $d = 1$. For $u \in \mathbb{R}$, let $M_t^u = \frac{e^{iuX_t}}{\varphi_t(u)}$, where $\varphi_t(u) = \mathbb{E}[e^{iuX_t}]$. Then M_t^u is a $\mathcal{G}_t$–martingale (cf. Propn. 3.1 and the remarks prior to the statement of this theorem). For $s_2 > s_1 > t$ and $u_1, u_2 \in \mathbb{R}$, we have

$$\begin{aligned}
\mathbb{E}[e^{i(u_1X_{s_1}+u_2X_{s_2})}|\mathcal{G}_{t+}] &= \lim_{w\downarrow t} \mathbb{E}[e^{i(u_1X_{s_1}+u_2X_{s_2})}|\mathcal{G}_w] \\
&= \lim_{w\downarrow t} \varphi_{s_2}(u_2) E[e^{iu_1X_{s_1}} M_{s_2}^{u_2}|\mathcal{G}_w] \\
&= \lim_{w\downarrow t} \varphi_{s_2}(u_2) E[e^{iu_1X_{s_1}} M_{s_1}^{u_2}|\mathcal{G}_w] \\
&= \varphi_{s_2-s_1}(u_2) \lim_{w\downarrow t} \mathbb{E}[e^{i(u_1+u_2)X_{s_1}}|\mathcal{G}_w]
\end{aligned}$$

(using the fact that the characteristic function of X_t is of the form $\varphi_t(u) = e^{t\eta(u)}$)

$$\begin{aligned}
&= \lim_{w\downarrow t} e^{i(u_1+u_2)X_w} \varphi_{s_1-w}(u_1+u_2)\varphi_{s_2-s_1}(u_2) \\
&= e^{i(u_1+u_2)X_t} \varphi_{s_1-t}(u_1+u_2)\varphi_{s_2-s_1}(u_2)
\end{aligned}$$

Similar reasoning shows that also

$$\mathbb{E}[e^{i(u_1X_{s_1}+u_2X_{s_2})}|\mathcal{G}_t] = e^{i(u_1+u_2)X_t} \varphi_{s_1-t}(u_1+u_2)\varphi_{s_2-s_1}(u_2)$$

and hence $\mathbb{E}[e^{i(u_1X_{s_1}+u_2X_{s_2})}|\mathcal{G}_{t+}] = \mathbb{E}[e^{i(u_1X_{s_1}+u_2X_{s_2})}|\mathcal{G}_t]$. We have now shown that $(*)$ holds. Now

$$\mathcal{M} = \{e^{i\sum_{j=1}^n \langle u_j, X_{s_j}\rangle} : n \in \mathbb{N}, s_1, \ldots, s_n \geq 0, u_1, \ldots, u_n \in \mathbb{R}^d\}$$

is a multiplicative class closed under conjugation. Let

$$\mathcal{H} = \{Z : Z \text{ a bounded RV with } \mathbb{E}[Z|\mathcal{G}_{t+}] = \mathbb{E}[Z|\mathcal{G}_t]\}$$

Then $\mathcal{H}$ is a vector space satisfying $\mathcal{H} \supseteq \mathcal{M}$. By a monotone class theorem, $\mathcal{H} \supseteq \mathrm{b}\sigma(\mathcal{M}) = \mathrm{b}\sigma(\bigcup_{s<\infty} \mathcal{F}_s^o)$, i.e. $\mathcal{H}$ contains every bounded $\sigma(\bigcup_{s<\infty} \mathcal{F}_s^o)$–measurable RV, i.e. $\mathbb{E}[Z|\mathcal{G}_{t+}] = \mathbb{E}[Z|\mathcal{G}_t]$ for every $Z \in \mathrm{b}\sigma(\bigcup_{s<\infty} \mathcal{F}_s^o)$. Since every event in $\mathcal{G}_t$ differs from an event in $\mathcal{F}_t^o$ by a null set, we can conclude $\mathcal{G}_t$ and $\mathcal{G}_{t+}$ differ at most by null sets. But both $\mathcal{G}_t, \mathcal{G}_{t+}$ include $\mathcal{N}$, and hence $\mathcal{G}_t = \mathcal{G}_{t+}$.

$\dashv$

D Lévy Processes without Jumps — *sans* Lévy Characterization

Here follows Bretagnolle's[**?**] direct proof of the fact that continuous Lévy processes are Brownian motions.

Proposition D.1 *If B_t is a one–dimensional centered Lévy process with continuous sample paths, then there is a $\sigma^2 \in \mathbb{R}^+$ such that*

$$\mathbb{E}[e^{iuB_t}] = e^{-t\sigma^2 u^2/2}$$

Hence B_t is a Brownian motion.

Proof: Since B_t has no jumps, it has moments of all orders, and thus its characteristic function is C^∞. Now recall that $\mathbb{E}[e^{iuB_t}] = e^{-t\psi(u)}$ for some C^∞ function ψ. Since $\psi'(0) = 0$ (because B_t is centered), we see — by differentiating repeatedly, and evaluating at $u = 0$ — that $\mathbb{E}[B_t^n] = \sum_{0<k<n} a_k t^k$ for some $a_k \in \mathbb{C}$. Now $\mathbb{E}B_t^2 = -t\psi''(0)$. By scaling, we may assume that $\mathbb{E}B_t^2 = t$. Now let $\pi = \{0 = t_0 < t_1 < \cdots < t_N = t\}$ be a partition of $[0,t]$. Then

$$\mathbb{E}[(\Delta_k B)^2] = \Delta_k t$$

where $\Delta_k B = B_{t_{k+1}} - B_{t_k}$, $\Delta_k t = t_{k+1} - t_k$. By Taylor's formula (second order) there exist $\theta \in [0,1]$ such that $e^{iux} = 1 + iux - \frac{1}{2}u^2 e^{iu\theta x}x^2 = 1 + iux - \frac{1}{2}u^2x^2 - \frac{1}{2}u^2x^2[e^{iu\theta x} - 1]$. Thus there exist $\theta_k \in [0,1]$ such that

$$e^{iuB_{t_{k+1}}} - e^{iuB_{t_k}} = e^{iuB_{t_k}}\left[iu\Delta_k B - \tfrac{1}{2}u^2(\Delta_k B)^2 - \tfrac{1}{2}u^2(\Delta_k B)^2[e^{iu\theta_k\Delta_k B} - 1]\right]$$

Define $\varphi_k(u) = \mathbb{E}[e^{iuB_{t_k}}]$. Then

$$\begin{aligned}
\mathbb{E}[e^{iuB_t} - 1] &= \sum_{0\le k<n} \mathbb{E}[e^{iuB_{t_{k+1}}} - e^{iuB_{t_k}}] \\
&= \sum_{0\le k<n} \mathbb{E}[e^{iuB_{t_k}}] \cdot \left(\mathbb{E}\left[iu\Delta_k B - \tfrac{1}{2}u^2(\Delta_k B)^2 - \tfrac{1}{2}u^2(\Delta_k B)^2[e^{iu\theta_k\Delta_k B} - 1]\right]\right) \\
&= 0 - \tfrac{1}{2}u^2 \sum_{0\le k<n} \varphi_k(u)\Delta_k t - \tfrac{1}{2}u^2 \sum_{0\le k<n} \varphi_k(u)\mathbb{E}\left[(\Delta_k B)^2[e^{iu\theta_k\Delta_k B} - 1]\right]
\end{aligned}$$

using the fact that increments are independent with mean zero and variance Δt. Now

$$\sum_{0\le k<n} \varphi_k(u)\Delta_k t \to \int_0^t e^{-s\psi(u)}\, ds \qquad \text{as mesh}(\pi)\to 0$$

Next, we investigate the third term, i.e.

$$\tfrac{1}{2}u^2 \sum_{0\le k<n} \varphi_k(u)\mathbb{E}\left[(\Delta_k B)^2[e^{iu\theta_k\Delta_k B}-1]\right]$$

By Taylor's Theorem, we have

$$e^{ix} = \sum_{k=0}^{n-1}\frac{(ix)^k}{k!} + \gamma\frac{|x|^n}{n!} \qquad \text{for some } \gamma\in\mathbb{C} \text{ with } |\gamma|\le 1$$

Thus

$$|e^{iu\theta_k\Delta_k B}-1| = \gamma_k|u||\Delta_k B| \qquad \text{for some } |\gamma_k|\le 1$$

using the fact that $|\theta_k|\le 1$ also. Now define, for $\alpha>0$, the event

$$A_\alpha = \{\sup_i \sup_{t_i\le r,s\le t_{i+1}} |B_r - B_s| < \alpha\}$$

Then on A_α we have $|e^{iu\theta_k\Delta_k B} - |\le \alpha|u|$, whereas on A_α^c (and, in fact, everywhere) we have $|e^{iu\theta_k\Delta_k B}-1|\le 2$. Hence we see that

$$\begin{aligned}
&\left|\tfrac{1}{2}u^2 \sum_{0\le k<n} \varphi_k(u)\mathbb{E}\left[(\Delta_k B)^2[e^{iu\theta_k\Delta_k B}-1]\right]\right| \\
&\le \tfrac{\alpha}{2}|u|^3\int_{A_\alpha}\sum_k(\Delta_k B)^2\, d\mathbb{P} + |u|^2\int_{A_\alpha^c}\sum_k(\Delta_k B)^2\, d\mathbb{P} \\
&\le \tfrac{\alpha}{2}|u^3|\ \mathbb{E}\Big[\sum_{0\le k<n}(\Delta_k B)^2\Big] + |u|^2\ \sqrt{\mathbb{P}(A_\alpha^c)}\cdot\sqrt{\mathbb{E}\Big[\Big(\sum_{0\le k<n}(\Delta_k B)^2\Big)^2\Big]} \\
&\qquad \text{by Hölder's } \mathcal{L}^2\text{–inequality} \\
&\le \tfrac{\alpha}{2}|u|^3 t + |u|^2\sqrt{\mathbb{P}(A_\alpha^c)}[\mathrm{O}(t+t^2+t^3)]^{\frac{1}{2}}
\end{aligned}$$

taking into account the condition on moments given above. Now as mesh$(\pi)\to 0$, we have $\mathbb{P}(A_\alpha^c)\to 0$, because the sample paths of B are a.s. continuous (and thus a.s. uniformly continuous on $[0,t]$). It follows that

$$\lim_{\text{mesh}(\pi)\to 0}\left|\tfrac{1}{2}u^2 \sum_{0\le k<n} \varphi_k(u)\mathbb{E}\left[(\Delta_k B)^2[e^{iu\theta_k\Delta_k B}-1]\right]\right| \le \tfrac{\alpha}{2}|u^3|t$$

Since $\alpha>0$ was arbitrary, this implies that

$$\lim_{\text{mesh}(\pi)\to 0}\tfrac{1}{2}u^2 \sum_{0\le k<n} \varphi_k(u)\mathbb{E}\left[(\Delta_k B)^2[e^{iu\theta_k\Delta_k B}-1]\right] = 0$$

Consequently,

$$e^{-t\psi(u)} - 1 = \mathbb{E}[e^{iuB_t}-1] = -\tfrac{1}{2}u^2\int_0^t e^{-s\psi(u)}\, ds$$

which implies that $\psi(u) = \frac{u^2}{2}$.

$\dashv$

E A Technical Result on Jumps of Martingales

We work on a stochastic base $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$, assumed to satisfy the usual conditions. All processes are assumed to be càdlàg and adapted.

Lemma E.1 *Suppose that $f, g : [0,t] \to \mathbb{R}$ are two càdlàg functions, null at zero, and that g is of finite variation on compacts. Let $\pi_n = \{0 = t_0^n < t_1^n < \cdots < t_{N(n)}^n = t\}$ be a sequence of partitions of $[0,t]$, with mesh$(\pi_n) \to 0$. Then*

$$\lim_{n\to\infty} \sum_{0\leq i<N(n)} \Big(f(t_{i+1}^n) - f(t_i^n)\Big) \cdot \Big(g(t_{i+1}^n) - g(t_i^n)\Big) = \sum_{s\leq t} \Delta f(s) \cdot \Delta g(s)$$

Proof: Define functions $t_+^n, t_-^n : [0,t] \to \mathbb{R}$ by

$$t_+^n = \sum_{0\leq i<N(n)} t_{i+1}^n I_{(t_i^n, t_{i+1}^n]} \qquad t_-^n = \sum_{0\leq i<N(n)} t_i^n I_{(t_i^n, t_{i+1}^n]}$$

Then

$$\begin{aligned}
&\sum_{0\leq i<N(n)} \Big(f(t_{i+1}^n) - f(t_i^n)\Big) \cdot \Big(g(t_{i+1}^n) - g(t_i^n)\Big) \\
&= \sum_{0\leq i<N(n)} \Big(f(t_{i+1}^n) - f(t_i^n)\Big) \int_{(t_i^n, t_{i+1}^n]} dg(s) \\
&= \int_{(0,t]} f(t_+^n(s)) - f(t_-^n(s))\, dg(s)
\end{aligned}$$

where the integrals make sense as Lebesgue–Stieltjes integrals, because g is of finite variation. Note that

$$t_+^n(s) \downarrow s \qquad t_-^n(s) \uparrow s \qquad \text{as } n \to \infty$$

and that $t_-^n(s) < s$ for all n. Because f is càdlàg, we have $f(t_+^n(s)) - f(t_-^n(s)) \to \Delta f(s)$. Thus, by the dominated convergence theorem

$$\lim_n \int_{(0,t]} f(t_+^n(s)) - f(t_-^n(s))\, dg(s) = \int_{(0,t]} \Delta f(s)\, dg(s) = \sum_{s\leq t} \Delta f(s) \cdot \Delta g(s)$$

$\dashv$

Proposition E.2 *Suppose that M, N are centered càdlàg martingales, that M is square–integrable, and that N has square–integrable variation (i.e. $\mathbb{E}[V_t(N)^2] < \infty$, where $V_t(N)$ is the variation of N on $[0,t]$). Then*

$$\mathbb{E}M_t N_t = \mathbb{E} \sum_{s\leq t} \Delta M_s \cdot \Delta N_s$$

Proof: Note that if $\pi = \{0 = t_0 < t_1 < \cdots < t_N = t\}$ is a partition of $[0,t]$, then

$$\mathbb{E}M_t N_t = \mathbb{E}\left(\sum_{0\leq k<N} (M_{t_{k+1}} - M_{t_k}) \sum_{0\leq k<N} (N_{t_{k+1}} - N_{t_k})\right) = \mathbb{E}\left[\sum_{0\leq k<N} (M_{t_{k+1}} - M_{t_k})(N_{t_{k+1}} - N_{t_k})\right]$$

by repeated applications of the tower property.

Now

$$\sum_{0\leq k<N}(M_{t_{k+1}}-M_{t_k})(N_{t_{k+1}}-N_{t_k})\leq 2\sup_{s\leq t}|M_s|\cdot V_t(N)$$

By Doob's $\mathcal{L}^2$–inequality, we have $\mathbb{E}[\sup_{s\leq t}|M_s|^2]\leq 4\mathbb{E}[M_t^2]<\infty$. Hence both $\sup_{s\leq t}|M_s|$ and $V_t(N)$ are in $\mathcal{L}^2$, and thus their product is in $\mathcal{L}^1$, i.e. $\sum_{0\leq k<N}(M_{t_{k+1}}-M_{t_k})(N_{t_{k+1}}-N_{t_k})$ is dominated by an integrable random variable. Now let $\pi_n=\{0=t_0^n<t_1^n<\cdots<t_{N(n)}^n=t\}$ be a sequence of partitions of $[0,t]$, with $\mathrm{mesh}(\pi_n)\to 0$. Then

$$\mathbb{E}M_tN_t=\mathbb{E}\left(\sum_{0\leq k<N(n)}(M_{t_{k+1}^n}-M_{t_k^n})(N_{t_{k+1}^n}-N_{t_k^n})\right)\qquad\text{for all } n\in\mathbb{N}$$

Therefore, by the dominated convergence theorem and the preceding lemma, we have

$$\mathbb{E}M_tN_t=\mathbb{E}\left[\lim_{n\to\infty}\sum_{0\leq k<N}(M_{t_{k+1}^n}-M_{t_k^n})(N_{t_{k+1}^n}-N_{t_k^n})\right]=\mathbb{E}\left[\sum_{s\leq t}\Delta M_s\cdot\Delta N_s\right]$$

$\dashv$

F Technical Results on Poisson Random Measures

Theorem F.1 *Suppose that $(E,\mathcal{B},\mu)$ is σ–finite measure space. Then there exists, on some probability space $(\Omega,\mathcal{F},\mathbb{P})$, a Poisson random measure M on $(E,\mathcal{B})$ with intensity measure μ.*

Proof: First assume that μ is a finite measure on $(E,\mathcal{B})$. If $\mu=0$, let $M=0$ identically. Else, let $(\Omega,\mathcal{F},\mathbb{P})$ be any probability space sufficiently rich to support the existence of a sequence $\{Z_n:n\in\mathbb{N}\}$ of random variables with common distribution $\frac{\mu}{\mu(E)}$, and a Poisson random variable Y with mean $\mu(E)$, such that that $Y,Z_1,Z_2,Z_3,\dots$ are mutually independent. Define, for $B\in\mathcal{B}$

$$M(B,\omega)=\sum_{k=1}^{Y(\omega)}I_B(Z_k(\omega))$$

where we define $\sum_1^0=0$. Then $M(B,\omega)$ counts the no. of Z_n's that belong to B, and is therefore a measure, for each ω. Note that $M(E)=Y$.

Now let $B_1,\dots,B_m\in\mathcal{B}$ be a partition of E, and let $n_1,\dots,n_m\in\mathbb{N}$ with $\sum_{k=1}^m n_k=n$. Then

$$\begin{aligned}
&\mathbb{P}(M(B_1)=n_2,\dots,M(B_m)=n_m)\\
&=\mathbb{P}(M(B_1)=n_2,\dots,M(B_m)=n_m|M(E)=n)\,\mathbb{P}(M(E)=n)\\
&=\mathbb{P}\left(\sum_{k=1}^n I_{B_1}(Z_k)=n_1,\dots,\sum_{k=1}^n I_{B_m}(Z_k)=n_k\right)\,\mathbb{P}(Y=n)\\
&=\frac{n!}{n_1!\dots n_m!}\left(\frac{\mu(B_1)}{\mu(E)}\right)^{n_1}\dots\left(\frac{\mu(B_m)}{\mu(E)}\right)^{n_m}\,e^{-\mu(E)}\frac{\mu(E)^n}{n!}\\
&=\prod_{k=1}^m e^{-\mu(B_k)}\frac{\mu(B_k)^{n_k}}{n_k!}
\end{aligned}$$

Summing over $n_1, \ldots, n_m$ except n_k, we obtain

$$\mathbb{P}(M(B_k) = n_k) = e^{-\mu(B_k)} \frac{\mu(B_k)^{n_k}}{n_k!}$$

This proves at once that the $M(B_k)$ are independent Poisson random variables with mean $\mu(B_k)$.

It remains to deal with the case where $\mu(E) = \infty$. Since μ is σ–finite, we can find disjoint $E_1, E_2, \dots \in \mathcal{B}$ such that $\bigcup_k E_k = E$ and such that each $\mu(E_k) < \infty$. We can then construct, on some probability space $(\Omega_k, \mathcal{F}_k, \mathbb{P}_k)$ a Poisson random measure M_k on $(E, \mathcal{B})$ with intensity measure $\mu|E_k$. By taking $(\Omega, \mathcal{F}, \mathbb{P})$ to be the product space of the $(\Omega_k, \mathcal{F}_k, \mathbb{P}_k)$, we can assume the M_k to be independent Poisson random measures $M_k : \mathcal{B} \times \Omega \to \mathbb{N} \cup \{\infty\}$ on $(E, \mathcal{B})$. Now define

$$M(B) = \sum_{k=1}^{\infty} M_k(B)$$

Then

$$\mathbb{E}[M(B)] = \sum_k \mathbb{E}[M_k(B)] = \sum_k \mu_k(B) = \mu(B)$$

Since a sum of independent Poisson variables is Poisson, we see that $M(B)$ is Poisson with mean $\mu(B)$ if $\mu(B) < \infty$. It remains to show that $M(B) = \infty$ a.s. if $\mu(B) = \infty$. First note that that there exists a constant a such that $(1 - e^{-r}) \geq \min\{\frac{r}{2}, a\}$: There is a constant b such that $(1 - e^{-r}) \geq \frac{r}{2}$ for $0 \leq r \leq b$. Now let $a = (1 - e^{-b})$. Now if $\mu(B) = \infty$, then

$$\sum_k \mathbb{P}(M_k(B) \geq 1) = \sum_k (1 - e^{-\mu_k(B)}) \geq \sum_k \min\{\tfrac{\mu_k(B)}{2}, a\} = \infty$$

Hence by the Borel–Cantelli lemma we see that $\mathbb{P}(M_k(B) \geq 1 \text{ i.o.}) = 1$, and thus that $\mathbb{P}(M(B) = \infty) = 1$ also.

$\dashv$

Lemma F.2 *Suppose that $N : \mathcal{B}(X) \times \Omega \to \bar{\mathbb{R}}_+$ is a random measure on a Borel subset X of $\mathbb{R}^d$, and suppose that the intensity μ of N (given by $\mu(B) = \mathbb{E}[N(B)]$ for $B \in \mathcal{B}(\mathbb{R}^d)$) is σ–finite. Suppose that*

(i) $N(B)$ is a Poisson random variable for each rectangle $B \in \mathcal{B}(\mathbb{R}^d)$*;*

(ii) If $B_1, B_2, \ldots, B_n$ are disjoint rectangles, then $N(B_1), N(B_2), \ldots, N(B_n)$ are independent.

Then N is a Poisson random measure with intensity μ.

Proof: (This proof seems unnecessarily complicated, and I'd like to find a simpler one.) For clarity of exposition, we assume $X = \mathbb{R}^d$. Suppose first that μ is a finite measure, and that U is a bounded open set. Then $U =\uparrow \bigcup_n C_n$ where each C_n is a finite union of disjoint rectangles, and the sequence $(C_n)_n$ is increasing. (E.g. for C_n, subdivide $\mathbb{R}^d$ into dyadic rectangles, with endpoints having coordinates of the form $k2^{-n}$. Let C_n be the union of all such subintervals which are contained in U.) Since each $C_n = \bigcup_{k \leq m} R_k$ is a finite union of disjoint rectangles, and since the $N(R_k)$ are, by hypothesis, independent Poisson random

variables with mean $\mu(R_k)$, we see that $N(C_n) = \sum_{k\leq m} N(R_k)$ a.s. is Poison with mean $\sum_{k\leq m}\mu(R_k) = \mu(C_n)$. Clearly $N(C_n) \uparrow N(U)$ a.s., by monotonicity of measures, and thus $N(C_n) \xrightarrow{w} N(U)$. In particular, comparing characteristic functions, we have

$$e^{\mu(C_n)[e^{it}-1]} \to \mathbb{E}[e^{itU}]$$

for all $t \in \mathbb{R}$, and since $\mu(C_n) \to \mu(U)$, we see that

$$\mathbb{E}[e^{itN(U)}] = e^{\mu(U)[e^{it}-1]}$$

Hence $N(U)$ is Poisson with mean $\mu(U)$.

Next suppose that $B \in \mathcal{B}(\mathbb{R}^d)$ is a bounded Borel set. Since every measure on a metric space is regular, there exists a decreasing sequence bounded of open sets (U_n) such that $B \subseteq U_n$ for each $n \in \mathbb{N}$, and such that $\mu(U_n) \downarrow \mu(B)$. Let $C = \bigcap_n U_n$. Then $B \subseteq C$ and $\mu(B) = \mu(C)$. Since $N(B) \leq N(C)$ a.s. and since $\mathbb{E}[N(B)] = \mathbb{E}[N(C)]$, we conclude that $N(B) = N(C)$ a.s. Now $N(U_n) \downarrow N(C)$ a.s., and thus in distribution. We have shown above that each $N(U_n)$ is Poisson with mean $\mu(U_n)$. Comparing characteristic functions (as above), we see that $N(C)$ is Poisson with mean $\lim_n \mu(U_n) = \mu(C)$. Since $N(B) = N(C)$ a.s., they are identically distributed, and so $N(B)$ is Poisson with mean $\mu(C) = \mu(B)$.

If $B \in \mathcal{B}(\mathbb{R}^d)$ is not necessarily bounded, and μ not necessarily finite, then we can write $B = \bigcup_n B_n$, where the B_n are disjoint bounded Borel sets, such that each $\mu(B_n)$ is finite. Then each $N(B_n)$ is Poisson with mean $\mu(B_n)$, and so $N(B) = \sum_n N(B_n)$ is Poisson with mean $\sum_n \mu(B_n) = \mu(B)$.

It remains to show that N has independent increments. If $U_1, \dots, U_m$ are disjoint open sets, then each can be written as a union of disjoint rectangles, $U_k = \bigcup_n C_{k,n}$. Hence $N(U_k) \in \sigma(N(C_{k,n}) : n \in \mathbb{N})$. These σ–algebras are independent, by hypothesis, and thus $N(U_1), \dots, N(U_m)$ are independent also.

In a metric space, every closed set is a countable intersection of open sets. Since a metric space is also a T_4–space, disjoint closed sets can be separated by disjoint open sets. Thus if $C_1, \dots, C_m$ are disjoint closed sets in $\mathbb{R}^d$, we can find sequences $(U_{k,n} : n \in \mathbb{N})$ of open sets such that $U_{k,n} \downarrow C_k$ (as $n \to \infty$), and such that $U_{k,n} \cap U_{j,l} = \emptyset$ is $j \neq k$. Since $N(C_k) \in \sigma(N(U_{k,n}) : n \in \mathbb{N})$, and since these σ–algebras are independent, we see that $N(C_1), \dots, N(C_m)$ are independent also.

Finally, if $B_1, \dots, B_m$ are disjoint Borel sets, then we may find sequences of closed sets $(C_{k,n} : n \in \mathbb{N})$ such that $C_{k,n} \subseteq B_k$ and $\mu(\bigcup_n C_{k,n}) = \mu(B_k)$, because μ is regular. Let $C_k = \bigcup_n C_{k,n}$. Then $C_k \subseteq B_k$ and $\mu(C_k) = \mu(B_k)$. As above, it follows that $N(B_k) = N(C_k)$ a.s. Now because $C_{k,n} \cap C_{j,l} = \emptyset$ when $j \neq k$, we see that $N(C_1), \dots, N(C_m)$ are independent, and thus $N(B_1), \dots N(B_m)$ are independent also.

$\dashv$

G Proof of the Lévy–Ito Decomposition Theorem

This section largely follows Bretagnolle[**?**]:

G.1 Preliminaries

Let $(X_t)_t$ be a $\mathbb{R}^d$–valued Lévy process on a stochastic base $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_t)$. Define

$$N_t(B) = \int_B N_t(dx) = \sum_{s \le t} I_B(\Delta X_s) \qquad \text{for } B \in \mathcal{B}(\mathbb{R}^d) \text{ bounded away from zero}$$
$$X_t(B) = \int_B x \; N_t(dx) = \sum_{s \le t} \Delta X_s I_B(\Delta X_s)$$

Let ν be the Lévy measure of X, i.e. ν is a measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ satisfying

$$\nu(B) = \mathbb{E}[N_1(B)] \qquad \text{and has} \quad \nu\{0\} = 0$$

Let J_X be the jump measure of X, i.e. $J_X : \mathcal{B}(H) \times \Omega \to \bar{\mathbb{Z}}_+$ is a random measure on $H = (0, \infty) \times \mathbb{R}^d \backslash \{0\}$ defined by

$$J_X(A) = \#\{s : (s, \Delta X_s) \in A\}$$

so that $J_X((0,t] \times B) = N_t(B)$. Then J_X is a Poisson random measure with intensity $\tilde{\nu} = \lambda \times \nu$.

It follows by Propn. 6.2 that for fixed t, N_t is a Poisson random measure on $\mathbb{R}^d \backslash \{0\}$ with intensity measure $t\nu$. By Propn. 6.8, each $X_t(B)$ is therefore a compound Poisson random variable with jump distribution $t\nu$. It is not hard to see that $X_t(B)$ is a compound Poisson process. Indeed:

Proposition G.1 *Suppose that X, N, ν are as above. Let $B \in \mathcal{B}(\mathbb{R}^d)$ be bounded away from zero, and let $f : \mathbb{R}^d \to \mathbb{R}^p$ be measurable. Define $Z_t = \int_B f(x) \; N_t(dx)$. Then Z_t is a compound Poisson process with intensity $\nu(B)$ and jump distribution $\sigma(\cdot) = \nu(B \cap f^{-1}(\cdot))$. Moreover, if $\int_B |f(x)|^2 \, d\nu < \infty$, then*

$$\mathbb{E}[Z_t] = t \int f(x) \; d\nu$$
$$\mathbb{E}\left[\left| Z_t - t \int_B f(x) \; d\nu \right|^2 \right] = t \int |f(x)|^2 \; d\nu$$

Proof: By Propn. 6.8, Z_t has a compound Poisson distribution with jump distribution $\sigma(A) = \nu(B \cap f^{-1}(A))$. It is clear that Z is càdlàg and piecewise constant, because $Z_t = \sum_{s \le t} f(\Delta X_s) I_B(\Delta X_s)$. To verify that Z is a Lévy process, we need only check that it has stationary independent increments. Now for $s < t$ we have $Z_t - Z_s = \sum_{s < u \le t} f(\Delta X_u) I_B(\Delta X_u) \in \sigma(X_v - X_u : s \le u < v \le t)$ which is independent of $\mathcal{F}_s$, because X has independent increments. By the stationarity of the increments of X, we conclude that $Z_t - Z_s$ has the same distribution as Z_{t-s}, so that Z also has stationary increments. By Thm. 4.7, Z is a compound Poisson process. Since Z jumps only when $N_t(B)$ jumps, the intensity (= expected no. of jumps per unit time) is $\nu(B)$. The remaining assertions follow by Propn. 6.8, using the fact that N_t is a Poisson random measure with intensity measure $t\nu$.

$\dashv$

Corollary G.2 *$X_t(B)$ is a d–dimensional compound Poisson process with intensity $\nu(B)$ and jump distribution $\nu(B \cap \cdot)$.*

The process $X_t - X_t(B)$ only has jumps with sizes in B^c. $X_t - X_t(B)$ is also a Lévy process. Indeed:

Proposition G.3 *Suppose that X, N, ν are as above. Let $B \in \mathcal{B}(\mathbb{R}^d)$ be bounded away from zero, and let $f : \mathbb{R}^d \to \mathbb{R}^p$ be measurable. Define*

$$Z_t = X_t - \int_B f(x)\ N_t(dx)$$

Then Z_t is a Lévy process.

Proof: We know that $X_t, \int_B f(x)\ N_t(dx)$ are Lévy processes. To check that Z_t is a Lévy process, we need only check that it has stationary independent increments. But for $s < t$ we have

$$Z_t - Z_s = X_t - X_s - \sum_{s<u\leq t} f(\Delta X_u) I_B(\Delta X_u)$$

which is $\sigma(X_v - X_u : s \leq u < v \leq t)$–measurable. Thus Z has independent increments. The stationarity of Z follows likewise from the stationarity of X.

$\dashv$

By removing the "large" jumps from X, we obtain a process with only "small" jumps. Suppose that we arbitrarily designate jumps of size $|\Delta X| \geq 1$ as "large". Then:

Corollary G.4 $X_t - \int_{|x|\geq 1} x\ N_t(dx)$ *is a Lévy process with moments of all orders.*

Proof: The preceding proposition shows that $X_t - \int_{|x|\geq 1} x\ N_t(dx)$ is a Lévy process. Since it only has jumps of size $|\Delta X| < 1$, it has moments of all orders, by Propn. 5.1.

$\dashv$

G.2 Martingale matters

Let $\mathcal{M}^2$ denote the space of square–integrable centered càdlàg martingales on $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_t)$ equipped with the topology induced by the seminorms $q_t(M) = \mathbb{E}[M_t^2]^{\frac{1}{2}}$. Note that By Doob's inequality

$$\mathbb{E}[\sup_{s\leq t} |M_s - N_s|^2] \leq 4 q_t(M - N)^2$$

Hence $\mathcal{M}^2$ is a metric space, with metric

$$d(M, N) = \sum_{n=1}^{\infty} 2^{-n} [q_n(M - N) \wedge 1]$$

Suppose now that $(M^n)_n$ is a Cauchy sequence in $\mathcal{M}^2$. For each t, we have $q_t(M^n - M^m) \to 0$ as $n, m \to 0$. Since $\mathcal{L}^2$–spaces are complete, there is a random variable M_t such that $M_t^n \to M_t$ in $\mathcal{L}^2$. It is easy to verify that $(M_t)_t$ is a martingale: Suppose $s \leq t$. Then

$$\begin{aligned} ||M_s - \mathbb{E}[M_t|\mathcal{F}_s]||_2 &\leq ||M_s - M_s^n||_2 + ||\mathbb{E}[M_t^n - M_t|\mathcal{F}_s]||_2 \\ &\leq ||M_s - M_s^n||_2 + ||M_t^n - M_t||^2 \end{aligned}$$

(using the fact that $|\mathbb{E}[X|\mathcal{F}_s]|^2 \le \mathbb{E}[|X|^2|\mathcal{F}_s]$.) Now the two terms on the righthand side can be made arbitrarily small by letting $n \to \infty$, and so $M_s = \mathbb{E}[M_t|\mathcal{F}_s]$. We can, of course, assume that M is càdlàg . (Indeed, by Doob's inequality

$$q_t(M^n - M) \to 0 \qquad \Rightarrow \qquad \mathbb{E}[\sup_{s\le t} |M^n_s - M_s|^2] \le 4q_t(M^n - M)^2 \to 0$$

and thus $M^n \to M$ uniformly in $\mathcal{L}^2$ on $[0,t]$. Hence we have a.s. uniform convergence along some fast subsequence $(M^{n_k})_k$, and this shows that the limit M is càdlàg also.) Since $\mathcal{L}^2$–convergence implies $\mathcal{L}^1$–convergence, we see that M is centered also. Hence:

Proposition G.5 *$\mathcal{M}^2$ is a complete locally convex metrizable space.*

□

Next, let X be a Lévy process with jump measure J_X and Lévy measure ν. Define $\tilde{J}_X$ by $\tilde{J}_X(d(t,x)) = J_X(d(t,x)) - \lambda(dt)\nu(dx)$. Similarly, define $\tilde{N}_t(dx) = N_t(dx) - t\nu(dx)$. Note that for each B bounded away from zero, the process $(\tilde{N}_t(B))_t$ is a martingale: $\tilde{N}_t(B)$ is clearly a Lévy process (because $N_t(B)$ is), and is moreover centered, because $\mathbb{E}[\tilde{N}_t(B)] = \mathbb{E}[N_t(B)] - t\nu(B) = 0$. Of course, this result can be extended: Define, for B bounded away from zero and $f \cdot I_B \in \mathcal{L}^2(\nu)$:

$$N_t(f,B) = \int_B f\; dN_t \qquad \tilde{N}_t(f,B) = \int_B f\; d\tilde{N}_t$$

Note that $N_t(B) = N_t(1,B)$, and that $N_t(f,B)$ is a compound Poisson process.

Proposition G.6 *Suppose that $B \in \mathcal{B}(\mathbb{R}^d)$ is bounded away from 0 and that $f : \mathbb{R}^d \to \mathbb{R}$ is such that $f \cdot I_B \in \mathcal{L}^2(\nu)$. Then $\tilde{N}_t(f,B)$ is a square–integrable centered càdlàg martingale.*

Proof: This follows directly from Propn. 6.8: N_t is a Poisson random measure with intensity measure $t\nu$, and hence $\mathbb{E}[\int_B f\; dN_t] = t\int_B f\; d\nu$, $\mathrm{Var}(\int_B f\; dN_t) = t\int_B |f|^2\; d\nu$.

⊣

Proposition G.7 *Suppose that B_1, B_2 are two disjoint Borel sets that are bounded away from zero, and that $f_1 \cdot I_{B_1}, f_2 \cdot I_{B_2} \in \mathcal{L}^2(\nu)$. Then the processes $N_t(f_1,B_1), N_t(f_2,B_2)$ are independent.*

Proof: For $u \in \mathbb{R}^d$ and $j = 1,2$, define $M^u_t(B_j) = \frac{e^{i\langle u, N_t(f_j,B_j)\rangle}}{\mathbb{E}[e^{i\langle u, N_t(f_j,B_j)\rangle}]} - 1$. Then each $M^u_t(B_j) \in \mathcal{M}^2$. Now each $N_t(f_j,B_j)$ is a compound Poisson process, from which it follows easily that the $M^u_t(B_j)$ have paths of finite variation on compacts. By Propn. 6.1, it follows that $\mathbb{E}[M^u_t(B_1)M^v_t(B_2)] = 0$ for all $u,v \in \mathbb{R}^d$, and thus that

$$\mathbb{E}[e^{i\langle u, N_t(f_1,B_1)\rangle} e^{i\langle u, N_t(f_2,B_2)\rangle} = \mathbb{E}[e^{i\langle u, N_t(f_1,B_1)\rangle}]\mathbb{E}[e^{i\langle u, N_t(f_2,B_2)\rangle}]$$

In particular, by independence and stationarity of increments, we see that

$$\begin{aligned}
&\mathbb{E}\left[e^{i\langle u_1, N_{t_1}(f_1,B_1)\rangle + i\sum_{k=2}^n \langle u_k, N_{t_k}(f_1B_1) - N_{t_{k-1}}(f_1,B_1)\rangle} e^{i\langle v_1, N_{t_1}(f_2,B_2)\rangle + i\sum_{k=2}^n \langle v_k, N_{t_k}(f_2,B_2) - N_{t_{k-1}}(f_2,B_2)\rangle}\right] \\
&= \mathbb{E}\left[e^{i\langle u_1, N_{t_1}(f_1,B_1)\rangle + i\sum_{k=2}^n \langle u_k, N_{t_k}(f_1,B_1) - N_{t_{k-1}}(f_1,B_1)\rangle}\right] \cdot \\
&\qquad \cdot \mathbb{E}\left[e^{i\langle v_1, N_{t_1}(f_2,B_2)\rangle + i\sum_{k=2}^n \langle v_k, N_{t_k}(f_2,B_2) - N_{t_{k-1}}(f_2,B_2)\rangle}\right]
\end{aligned}$$

from which follows the independence of $N_t(f_1,B_1)$ and $N_t(f_2,B_2)$.

⊣

G.3 Decomposition for centered Lévy processes with jumps bounded by 1

In this subsection, we prove the Lévy–Ito decomposition theorem for d–dimensional Lévy processes that are centered (hence martingales) and have jump amplitudes ≤ 1.

Let X be such a Lévy process, with Lévy measure ν. Note that X has moments of all orders, and thus that $X \in \mathcal{M}^2$. Now define $M_t(B) = \tilde{N}_t(\mathrm{id}, B) = X_t(B) - t\int_B x\ \nu(dx) = \int_B x\ [N_t(dx) - \nu(dx)]$. For $n \in \mathbb{N}$, define $D_n = \{x \in \mathbb{R}^d : \frac{1}{n+1} < |x| < \frac{1}{n}\}$, $C_n = \bigcup_{k\leq n} D_k$. Then the $M_t(D_n)$ are pairwise independent d–dimensional martingales/Lévy processes, by Propn. G.7. Moreover, the component processes $X_t^{(j)} - M_t^{(j)}(C_n)$ and $M_t^{(j)}(C_n)$ independent for $j = 1, \dots, d$, as can easily be verified by an argument akin to the proof of Propn. G.7. It follows that

$$\mathrm{Var}(X_t^{(j)}) = \mathrm{Var}(X_t^{(j)} - M_t^{(j)}(C_n)) + \mathrm{Var}(M_t^{(j)}(C_n))$$

and thus that

$$\mathrm{Var}(M_t^{(j)}(C_n)) \leq \mathrm{Var}(X_t^{(j)}) < \infty$$

for all $n \in \mathbb{N}$. By Propn. 6.8, $\mathrm{Var}(M_t^{(j)}(C_n)) = t\int_{C_n} x_j^2\ \nu(dx)$, where x_j is the j^{th} component of $x \in \mathbb{R}^d$. Thus $\mathrm{Var}(M_t^{(j)}(C_n))$ is increasing (as $n \to \infty$), and bounded by $\mathrm{Var}(X_t^{(j)})$, hence convergent, and therefore a Cauchy sequence. Now if $n < m$, then

$$\mathbb{E}\left[\left(M_t^{(j)}(C_m) - M_t^{(j)}(C_n)\right)^2\right] = t\int_{C_m} x_j^2\ \nu(dx) - t\int_{C_n} x_j^2\ \nu(dx)$$

so $(M_t^{(j)}(C_n))_{t\geq 0}$ is a Cauchy sequence in $\mathcal{M}^2$, and hence converges. It follows that the sequence $M_t(C_n)$ of d–dimensional martingales converges also. Note that

$$C_n \uparrow \{x \in \mathbb{R}^d : |x| \leq 1\}\backslash\{0\} \qquad \text{as } n \to \infty$$

and that $\nu\{0\} = 0$.

Define X_t^d to be the limit of the $M_t(C_n)$ in $\mathcal{M}^2$ (as $n \to \infty$), i.e. $X_t^d = \int_{|x|\leq 1} x\tilde{N}_t(dx)$. Further define $X_t^c = X_t - X_t^d$. Clearly, X_t^c is the $\mathcal{M}^2$–limit of the sequence $X_t - M_t(C_n)$. Since $X_t - M_t(C_n)$ and $M_t(C_n)$ are independent, so are their limits, i.e. X_t^d, X_t^c are independent.

We now verify that X_t^c is continuous: Indeed, by Doob's inequality, the $\mathcal{L}^2$–convergence of the $X_t - M_t(C_n)$ implies uniform convergence on compacts in $\mathcal{L}^2$ (specifically, $\sup_{s\leq t} |X_t - M_t(C_n) - X_t^c|^2 \to 0 \in \mathcal{L}^2$), and thus a.s. uniform convergence along a fast subsequence. Since $X_t - M_t(C_n)$ only has jumps of size $\leq \frac{1}{n+1}$, uniform convergence implies that X_t^c has no jumps at all.

Next, we show that X^c, X^d are Lévy processes. Indeed, X_t^d is the $\mathcal{L}^2$–limit of $M_t(C_n)$. Thus, by the dominated convergence theorem and the fact that the $M_t(C_n)$ are Lévy processes, we have $\mathbb{E}[e^{i\langle u, X_t^d - X_s^d\rangle}|\mathcal{F}_s] = \lim_n \varphi_{M_{t-s}(C_n)}(u) = \mathbb{E}[e^{i\langle u, X_{t-s}^d\rangle}]$ (where φ_Z denotes the characteristic function of the random variable Z). This proves that $X_t^d - X_s^d$ has independent identically distributed increments, and thus that X_t^d is a Lévy process. The same argument shows that X_t^c is a Lévy process.

Arguing in exactly the same way as in Propn. G.7, we can show that X_t^c is independent of $N_t(f, B)$ whenever f is bounded away from zero and $f \cdot I_B \in \mathcal{L}^2(\nu)$: Define $M_t^u = \frac{e^{i\langle u, X_t^c\rangle}}{\mathbb{E}[e^{i\langle u, X_t^c\rangle}]} - 1$ and $N_t^v = \frac{e^{i\langle v, N_t(f,B)\rangle}}{\mathbb{E}[e^{i\langle v, N_t(f,B)\rangle}]} - 1$. Then M_t^u is a continuous martingale, and hence $\mathbb{E}[M_t^u N_t^v] = 0$, by Propn. 6.1. The rest follows as Propn. G.7.

We have shown:

Theorem G.8 *Suppose that X_t is a centered Lévy process whose jumps are bounded by 1. Then we have*

$$X_t = X_t^c + X_t^d$$

where X_t^c is a martingale with continuous sample paths,

$$X_t^d = \int_{|x|\leq 1} x \, [N_t(dx) - t\nu(dx)]$$

is a martingale, and X_t^c, X_t^d are independent Lévy processes. Moreover, the Poisson random measure N_t is independent of X_t^c.

□

Remarks G.9 (a) Of course, there is nothing special about the constant 1 in the preceding theorem. We could have proved a similar decomposition theorem for centered Lévy processes whose jump sizes are bounded by any positive constant.

(b) Note that the integral

$$\int_{|x|\leq 1} x N_t(dx)$$

may well diverge. So, indeed, may the integral $\int_{|x|\leq 1} x \ \nu(dx)$. It is the *compensated* integral $\int_{|x|\leq 1} x \ \tilde{N}_t(dx) = \int_{|x|\leq 1} x \ [N_t(dx) - t\nu(dx)]$ which is guaranteed to be finite.

(c) Since $M_t(C_n) \to X_t^d$ in $\mathcal{L}^2$, we see that

$$\infty > \mathbb{E}[|X_t^d|^2] = \lim_n \mathbb{E}[|M_t(C_n)|^2] = t \int_{|x|\leq 1} |x|^2 \ \nu(dx)$$

Though $\int_{|x|\leq 1} x \ N_t(dx)$ may not converge, the integral $\int_{|x|\leq 1} |x|^2 \ \nu(dx)$ is necessarily finite. Moreover, because $\nu(B) < \infty$ when B is bounded away from zero, we must have $\int_{|x|>1} \ \nu(dx) < \infty$. It follows that ν is a Borel measure with the properties that

$$\int_{\mathbb{R}^d} |x|^2 \wedge 1 \ \nu(dx) < \infty \qquad \nu\{0\} = 0$$

A measure with these properties is sometimes called a *Lévy measure*. Thus, the Lévy measure of a Lévy process is a Lévy measure(!!)

□

G.4 The General Case

Suppose that X_t is a d–dimensional Lévy process. Define the Poisson random measure N_t on $\mathcal{B}(\mathbb{R}^d \backslash \{0\})$ by $N_t(B) := \sum_{s\leq t} I_B(\Delta X_s)$, and let $\nu(B) = \mathbb{E}[N_1(B)]$ be the Lévy measure of X. Recall the convention that $\nu\{0\} = 0$.

By Propn. G.3, the process $X_t - \int_{|x|>1} x \ N_t(dx)$ is a Lévy process, with jumps bounded by 1. It therefore has moments of all orders, by Propn. 5.1. Define $\gamma \in \mathbb{R}^d$ by $\gamma t = \mathbb{E}[X_t - \int_{|x|>1} x \ N_t(dx)]$, so that $\tilde{X}_t := X_t - \int_{|x|>1} x \ N_t(dx) - \gamma t$ is a centered Lévy process

with jumps bounded by 1. Note that $\tilde{X}_t$ and X_t have the same jumps with sizes ≤ 1 (i.e. if $|\Delta X_t| \leq 1$, then $\Delta \tilde{X}_t = \Delta X_t$) and thus that the Lévy measure of $\tilde{X}_t$, when restricted to Borel subsets of $\{x : 0 < |x| \leq 1\}$, is equal to ν. By Thm. G.8, we have the decomposition $\tilde{X}_t = \tilde{X}_t^c + \tilde{X}_t^d$, where $\tilde{X}^c, \tilde{X}^d$ are independent centered Lévy processes (hence martingales), $\tilde{X}_t^d = \int_{|x|\leq 1} x\, [N_t(dx) - t\nu(dx)]$ and $\tilde{X}_t^c$ has continuous sample paths. Also, $\tilde{X}^c$ and $N_t(B)$ are independent whenever B is bounded away from zero.

It follows by Theorem 5.3 that $\tilde{X}_t^c$ is a Brownian motion with mean 0 and covariance matrix A. We therefore have

$$X_t = \tilde{X}_t^c + \int_{|x|\leq 1} x\, [N_t(dx) - t\nu(dx)] + \int_{|x|>1} x\, N_t(dx) + \gamma t$$

We have proved:

Theorem G.10 (Lévy–Ito Decomposition)*Let X be a d–dimensional Lévy process. Then X has decomposition*

$$X_t = \gamma t + B_t + \int_{|x|\leq 1} x\, [N_t(dx) - t\nu(dx)] + \int_{|x|>1} x\, N_t(dx)$$

Here, $\gamma := \mathbb{E}[X_t - \int_{|x|>1} x\, N_t(dx)] \in \mathbb{R}^d$.

B_t is a centered Brownian motion with covariance matrix A.

The process $\int_{|x|\leq 1} x\, [N_t(dx) - t\nu(dx)]$ is a martingale independent of B_t.

For each $B \in \mathcal{B}(\mathbb{R}^d)$ bounded away from zero and each $f \cdot I_B \in \mathcal{L}^2(\nu)$, B_t is independent of $\int_B f(x)\, N_t(dx)$.

Furthermore, $\int_{B_1} f_1\, N_t(dx)$ and $\int_{B_2} f_2\, N_t(dx)$ are independent whenever B_1, B_2 are disjoint Borel sets which are bounded away from zero (assuming $f_1 \cdot I_{B_1}, f_2 \cdot I_{B_2} \in \mathcal{L}^2(\nu)$).

The measure ν is a Lévy measure, i.e. satisfies $\int_{\mathbb{R}^d} |x|^2 \wedge 1\, \nu(dx) < \infty$, and $\nu\{0\} = 0$.

$\square$